

# AN ANALYTICAL FRAMEWORK FOR A CONSENSUS-BASED GLOBAL OPTIMIZATION METHOD

JOSÉ A. CARRILLO, YOUNG-PIL CHOI, CLAUDIA TOTZECK, AND OLIVER TSE

**ABSTRACT.** In this paper we provide an analytical framework for investigating the efficiency of a consensus-based model for tackling global optimization problems. This work should shed light on the prospects of justifying the efficacy of the optimization algorithm in the mean-field sense. Extensions of the mean-field equation to include nonlinear diffusion of porous medium type is introduced. Theoretical results on decay estimates are then underlined by numerical simulations.

**Keywords:** Global optimization, opinion dynamics, consensus formation, agent-based models, stochastic dynamics, mean-field limit

## 1. INTRODUCTION

Over the last decades, individual-based models have been widely used in the investigation of complex systems that manifest self-organization or collective behavior. Examples of such complex systems include swarming behavior, crowd dynamics, opinion formation, synchronization, and many more, that are present in the field of mathematical biology, ecology and social dynamics, see for instance [5, 6, 13, 14, 18, 22, 26, 28, 34] and the references therein.

In the field of global optimization, *metaheuristics*, such as evolutionary algorithms [4, 1, 32] and swarm intelligence [20, 24], have played an increasing role in the design of fast algorithms to provide sufficiently good solutions in tackling hard optimization problems, which includes the traveling salesman problem that is known to be NP hard. Metaheuristics, in general, may be considered as high level concepts for exploring search spaces by using different strategies, chosen in such a way, that a dynamic balance is achieved between the exploitation of the accumulated search experience and the exploration of the search space [9]. Notable metaheuristics for global optimization include, for example, the Ant Colony Optimization, Genetic Algorithms, Particle Swarm Optimization and Simulated Annealing, all of which are stochastic in nature [7]. Unfortunately, despite having to stand the test of time, a majority of metaheuristic methods lack the proper justification of its efficacy in the mathematical sense. The universal intent of research in the field is to ascertain whether a given metaheuristic is capable of finding an optimal solution when provided with sufficient information. Due to the stochastic nature of metaheuristics, answers to this question are non-trivial, and they are always probabilistic.

Recently, the use of opinion dynamics and consensus formation in global optimization has been introduced in [31], where the authors showed substantial numerical and partial analytical evidence of its applicability to solving multi-dimensional optimization problems of the form

$$\min_{x \in \Omega} f(x), \quad \Omega \subset \mathbb{R}^d \text{ a domain,}$$

for a given cost function  $f \in \mathcal{C}_b(\mathbb{R}^d)$ , which is assumed to be non-negative without loss of generality.

The optimization algorithm involves the use of multiple agents located within the domain  $\Omega$  to dynamically establish a consensual opinion amongst themselves in finding the global minimizer to the minimization problem, while taking into consideration the opinion of all active agents. First order models for consensus have been studied in the mathematical community interested in granular materials and swarming leading to aggregation-diffusion and kinetic equations, which have nontrivial stationary states or flock solutions, see [11, 16, 17, 12] and the references therein.

They are also common tools in control engineering to establish consensus in graphs [29, 37] for instance among many others.

In order to achieve the goal of optimizing a given function  $f(x)$ , we consider an interacting stochastic system of  $N \in \mathbb{N}$  agents with position  $X_t^i \in \mathbb{R}^d$ , described by the system of stochastic differential equations

$$(1a) \quad dX_t^i = -\lambda(X_t^i - v_t) dt + \sigma |X_t^i - v_t| dW_t^i,$$

$$(1b) \quad v_t = \sum_{i=1}^N X_t^i \left( \frac{\omega_f^\alpha(X_t^i)}{\sum_{i=1}^N \omega_f^\alpha(X_t^i)} \right),$$

with  $\lambda, \sigma > 0$ , where  $\omega_f^\alpha$  is a weight, which we take as  $\omega_f^\alpha(x) = \exp(-\alpha f(x))$  for some appropriately chosen  $\alpha > 0$ . Notice that (1) resembles a geometric Brownian motion, which drifts in the direction  $v_t \in \mathbb{R}^d$ . This system is a simplified version of the algorithm introduced in [31], while keeping the essential ingredients and mathematical difficulties. The first term in (1) imposes a global relaxation towards a position determined by the behavior of the normalized moment given by  $v_t$ , while the diffusion term tries to concentrate again around the behavior of  $v_t$ . In fact, agents with a position differing a lot from  $v_t$  are diffused more and then they explore a larger portion of the landscape of the graph of  $f(x)$ , while the explorer agents closer to  $v_t$  diffuse much less. The normalized moment  $v_t$  is expected to dynamically approach the global minimum of the function  $f$ , at least when  $\alpha$  is large enough, see below. This idea is also used in simulated annealing algorithms. The well-posedness of this system will be thoroughly investigated in Section 2.

Formal passage to the mean-field limit,  $N \rightarrow \infty$ , for this system yields the *nonlinear process*

$$(2a) \quad d\bar{X}_t = -\lambda(\bar{X}_t - v_f[\rho_t]) dt + \sigma |\bar{X}_t - v_f[\rho_t]| dW_t$$

$$(2b) \quad v_f[\rho_t] = \int x d\eta_t^\alpha, \quad \eta_t^\alpha = \omega_f^\alpha \rho_t / \|\omega_f^\alpha\|_{L^1(\rho_t)}, \quad \rho_t = \text{law}(\bar{X}_t),$$

subject to the initial condition  $\text{law}(\bar{X}_0) = \rho_0$ . We call  $\eta_t^\alpha$  the  $\alpha$ -weighted measure.

The measure  $\rho_t = \text{law}(\bar{X}_t) \in \mathcal{P}(\mathbb{R}^d)$  is a Borel probability measure, which describes the evolution of a one-particle mean-field distribution. In the case that  $\rho_t$  is absolutely continuous w.r.t. the Lebesgue measure  $dx$ , i.e.,  $d\rho_t = u_t dx$ , for some  $u_t \in L^1(\mathbb{R}^d; dx)$ , we recall from [31, Remark 1], see also [19], that  $\omega_f^\alpha \rho_t$  satisfies the well-known *Laplace principle*:

$$\lim_{\alpha \rightarrow \infty} \left( -\frac{1}{\alpha} \log \left( \int e^{-\alpha f} d\rho_t \right) \right) = \inf f \geq 0.$$

Therefore, if  $f$  attains a single minimum  $x_* \in \text{supp}(\rho_t)$ , then the  $\alpha$ -weighted measure  $\eta_t^\alpha \in \mathcal{P}(\mathbb{R}^d)$  approximates a Dirac distribution  $\delta_{x_*}$  at  $x_* \in \mathbb{R}^d$  for large  $\alpha \gg 1$ . Consequently, the first moment of  $\eta_t^\alpha$ , given by  $v_f[\rho_t]$ , provides a good estimate of  $x_* = \arg \min f$ .

The (infinitesimal) generator corresponding to the nonlinear process (2a) is given by

$$(3) \quad L\varphi = \kappa \Delta \varphi - \mu \cdot \nabla \varphi, \quad \text{for } \varphi \in \mathcal{C}_c^2(\mathbb{R}^d),$$

with drift and diffusion coefficients

$$\mu_t = \lambda(x - v_f[\rho_t]), \quad \kappa_t = (\sigma^2/2)|x - v_f[\rho_t]|^2,$$

respectively. Therefore, the Fokker–Planck equation reads

$$(4) \quad \partial_t \rho_t = \Delta(\kappa \rho_t) + \nabla \cdot (\mu \rho_t), \quad \lim_{t \rightarrow 0} \rho_t = \rho_0,$$

where  $\rho_t \in \mathcal{P}(\mathbb{R}^d)$  for  $t \geq 0$  satisfies (4) in the weak sense.

Notice that the Fokker–Planck equation (4) is a nonlocal, nonlinear degenerate drift-diffusion equation, which makes its analysis a nontrivial task. Its well-posedness will be the topic of Section 3. Furthermore, using established ideas taken from [10, 33], we make the passage to mean-field rigorous in Section 4. Having justified the validity of the Fokker–Planck equation (4) as a mean-field limit of the microscopic system (1), we proceed to give justifications for the applicability of the microscopic system (1) as a tool for solving global optimization problems, via its mean-field counterpart.

More specifically, we will show in Section 5 that under certain assumptions on the cost function  $f$ , one obtains a uniform consensus as the limiting measure ( $t \rightarrow \infty$ ) corresponding to (4), i.e.,

$$\rho_t \longrightarrow \delta_{\hat{x}} \quad \text{as } t \rightarrow \infty,$$

for some  $\hat{x} \in \mathbb{R}^d$  possibly depending on the initial density  $\rho_0$ . It is also shown that this convergence happens exponentially in time. Moreover, under the same assumptions on  $f$ , the point of consensus  $\hat{x}$  may be made sufficiently close to  $x_* = \arg \min f$  by choosing  $\alpha \gg 1$  sufficiently large, which is the main goal for global optimization.

We conclude the paper with an extension of the Fokker–Planck equation (4) to include nonlinear diffusion of porous medium type and provide numerical evidence for consensus formation in the one dimensional case. For this reason, we introduce an equivalent formulation of the mean-field equation in terms of the pseudo-inverse distribution  $\chi_t(\eta) = \inf\{x \in \mathbb{R} \mid \rho_t((-\infty, x]) > \eta\}$ . We also compare the microscopic approximation corresponding to the porous medium type Fokker–Planck equation with the original consensus-based microscopic system (1) and the proposed algorithm in [31], showcasing the exponential decay rate of the error in suitable transport distances towards the global minimizers.

## 2. WELL-POSEDNESS OF THE MICROSCOPIC MODEL

In this section we study the existence of a unique process  $\{(X_t^{(1,N)}, \dots, X_t^{(N,N)}) \mid t \geq 0\}$ , which satisfies our consensus-based optimization scheme (1). We denote  $\mathbf{X}_t^{(N)} := (X_t^{(1,N)}, \dots, X_t^{(N,N)})^\top$ , and write, for an arbitrary but fixed  $N \in \mathbb{N}$ , system (1) as

$$(5) \quad d\mathbf{X}_t^{(N)} = -\lambda \mathbf{F}_N(\mathbf{X}_t^{(N)}) dt + \sigma \mathbf{M}_N(\mathbf{X}_t^{(N)}) d\mathbf{W}_t^{(N)},$$

where  $\mathbf{W}_t := (W_t^{(1,N)}, \dots, W_t^{(N,N)})^\top$  denotes the standard Wiener process in  $\mathbb{R}^{Nd}$ , and

$$\begin{aligned} \mathbf{F}_N(\mathbf{X}) &= (F_N^1(\mathbf{X}), \dots, F_N^N(\mathbf{X}))^\top \in \mathbb{R}^{Nd}, \quad F_N^i(\mathbf{X}) = \frac{\sum_{j \neq i} (X^i - X^j) \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \in \mathbb{R}^d, \\ \mathbf{M}_N(\mathbf{X}) &= \text{diag}(|F_N^1(\mathbf{X})| \mathbb{I}_d, \dots, |F_N^N(\mathbf{X})| \mathbb{I}_d) \in \mathbb{R}^{Nd \times Nd}. \end{aligned}$$

At this point, we make smoothness assumptions regarding our cost function  $f$ .

**Assumption 1.** The cost function  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  is locally Lipschitz continuous and nonnegative. In short,  $f \in \text{Lip}_{loc}(\mathbb{R}^d)$ ,  $f \geq 0$ .

Under these conditions on  $f$ , we easily deduce that  $F_N^i$ ,  $1 \leq i \leq N$ , is locally Lipschitz continuous and has linear growth. Consequently,  $\mathbf{F}_N$  and  $\mathbf{M}_N$  are locally Lipschitz continuous and have linear growth. To be more precise, we obtain the following result.

**Lemma 2.1.** Let  $N \in \mathbb{N}$ ,  $\alpha, k > 0$  be arbitrary. Then for any  $\mathbf{X}, \hat{\mathbf{X}} \in \mathbb{R}^{Nd}$  with  $|\mathbf{X}|, |\hat{\mathbf{X}}| \leq k$ :

$$\begin{aligned} |F_N^i(\mathbf{X}) - F_N^i(\hat{\mathbf{X}})| &\leq |X^i - \hat{X}^i| + \left(1 + \frac{2c_k}{N} \sqrt{N|\hat{X}^i|^2 + |\hat{\mathbf{X}}|^2}\right) |\mathbf{X} - \hat{\mathbf{X}}|, \\ |F_N^i(\mathbf{X})| &\leq |X^i| + |\mathbf{X}|, \end{aligned}$$

with constant  $c_k = \alpha \|\nabla f\|_{L^\infty(B_k)} \exp(\alpha \|f\|_{L^\infty(B_k)})$ , where  $B_k = \{x \in \mathbb{R}^d \mid |x| \leq k\}$ .

*Proof.* Let  $\mathbf{X}, \hat{\mathbf{X}} \in \mathbb{R}^{Nd}$  with  $|\mathbf{X}|, |\hat{\mathbf{X}}| \leq k$  for some  $k \geq 0$ . Then

$$F_N^i(\mathbf{X}) - F_N^i(\hat{\mathbf{X}}) = \frac{\sum_{j \neq i} (X^i - X^j) \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} - \frac{\sum_{j \neq i} (\hat{X}^i - \hat{X}^j) \omega_f^\alpha(\hat{X}^j)}{\sum_j \omega_f^\alpha(\hat{X}^j)} = \sum_{\ell=1}^3 I_\ell,$$

where the terms  $I_\ell$ ,  $\ell = 1, 2, 3$ , are given by

$$\begin{aligned} I_1 &= \frac{\sum_{j \neq i} (X^i - \hat{X}^i + \hat{X}^j - X^j) \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)}, \quad I_2 = \frac{\sum_{j \neq i} (\hat{X}^i - \hat{X}^j) (\omega_f^\alpha(X^j) - \omega_f^\alpha(\hat{X}^j))}{\sum_j \omega_f^\alpha(X^j)}, \\ I_3 &= \sum_{j \neq i} (\hat{X}^i - \hat{X}^j) \omega_f^\alpha(\hat{X}^j) \frac{\sum_j (\omega_f^\alpha(\hat{X}^j) - \omega_f^\alpha(X^j))}{\sum_j \omega_f^\alpha(X^j) \sum_j \omega_f^\alpha(\hat{X}^j)}, \end{aligned}$$

that may easily be estimated by

$$\begin{aligned} |I_1| &\leq |X^i - \hat{X}^i| + |\mathbf{X} - \hat{\mathbf{X}}|, \quad |I_2| \leq \frac{c_k}{N} |\mathbf{X} - \hat{\mathbf{X}}| \sqrt{N|\hat{X}^i|^2 + |\hat{\mathbf{X}}|^2}, \\ |I_3| &\leq \frac{c_k}{N} |\mathbf{X} - \hat{\mathbf{X}}| \sqrt{N|\hat{X}^i|^2 + |\hat{\mathbf{X}}|^2}. \end{aligned}$$

Putting all these terms together yields the required estimate.

As for the estimate of  $|F_N^i(\mathbf{X})|$ , we easily obtain

$$|F_N^i(\mathbf{X})| = \left| X^i - \frac{\sum_j X^j \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \right| \leq |X^i| + |\mathbf{X}|,$$

thereby concluding the result.  $\square$

Due to Lemma 2.1, we may invoke standard existence results of strong solutions for (5) [21].

**Theorem 2.1.** *For each  $N \in \mathbb{N}$ , the stochastic differential equation (5) has a unique strong solution  $\{\mathbf{X}_t^{(N)} \mid t \geq 0\}$  for any initial condition  $\mathbf{X}_0^{(N)}$  satisfying  $\mathbb{E}|\mathbf{X}_0^{(N)}|^2 < \infty$ .*

*Proof.* As mentioned above, we make use of a standard result on existence of a unique strong solution. To this end, we show the existence of a constant  $b_N > 0$ , such that

$$(6) \quad -2\lambda \mathbf{X} \cdot \mathbf{F}_N(\mathbf{X}) + \sigma^2 \text{trace}(\mathbf{M}_N \mathbf{M}_N^\top)(\mathbf{X}) \leq b_N |\mathbf{X}|^2.$$

Indeed, since the following inequalities hold:

$$\begin{aligned} -X^i \cdot F_N^i(\mathbf{X}) &= -X^i \cdot \frac{\sum_{j \neq i} (X^i - X^j) \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \leq -|X^i|^2 + |X^i| |\mathbf{X}|, \\ |F_N^i(\mathbf{X})|^2 &= \left| \frac{\sum_{j \neq i} (X^i - X^j) \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \right|^2 \leq 2(|X^i|^2 + |\mathbf{X}|^2), \end{aligned}$$

we conclude that

$$\begin{aligned} -2\lambda \langle \mathbf{X}, \mathbf{F}_N(\mathbf{X}) \rangle + \sigma^2 \text{trace}(\mathbf{M}_N \mathbf{M}_N^\top)(\mathbf{X}) &= \sum_i (-2\lambda \langle X^i, F_N^i(\mathbf{X}) \rangle + d\sigma^2 |F_N^i(\mathbf{X})|^2) \\ &\leq \sum_i 2\lambda (-|X^i|^2 + |X^i| |\mathbf{X}|) + 2d\sigma^2 (|X^i|^2 + |\mathbf{X}|^2) \\ &\leq 2(\lambda\sqrt{N} + 2d\sigma^2 N) |\mathbf{X}|^2 =: b_N |\mathbf{X}|^2. \end{aligned}$$

Along with the local Lipschitz continuity and linear growth of  $\mathbf{F}_N$  and  $\mathbf{M}_N$ , we obtain the assertion by applying [21, Theorem 3.1].  $\square$

*Remark 2.1.* In fact, the estimate (6) yields a uniform bound on the second moment of  $\mathbf{X}_t$ . Indeed, by application of the Itô formula, we obtain

$$\frac{d}{dt} \mathbb{E}|\mathbf{X}_t^{(N)}|^2 = -2\lambda \mathbb{E}[\langle \mathbf{X}_t, \mathbf{F}_N(\mathbf{X}_t^{(N)}) \rangle] + \sigma^2 \mathbb{E}[\text{trace}(\mathbf{M}_N \mathbf{M}_N^\top)(\mathbf{X}_t^{(N)})] \leq b_N \mathbb{E}|\mathbf{X}_t^{(N)}|^2.$$

Therefore, the Gronwall inequality yields

$$\mathbb{E}|\mathbf{X}_t^{(N)}|^2 \leq e^{b_N t} \mathbb{E}|\mathbf{X}_0^{(N)}|^2 \quad \text{for all } t \geq 0,$$

i.e., the solution exists globally in time for each fixed  $N \in \mathbb{N}$ .

Unfortunately, for the mean-field limit ( $N \rightarrow \infty$ ) we lose control of the previous bound, since  $b_N \rightarrow \infty$  as  $N \rightarrow \infty$ . Therefore, we will need a finer moment estimate on  $\mathbf{X}^{(N)}$ . In particular, we need an  $N$  independent moment estimate for  $\mathbf{X}^{(N)}$ .

Consider the evolution of  $\xi_t^{ij} := X_t^{(i,N)} - X_t^{(j,N)}$  given by

$$d\xi_t^{ij} = -\lambda \xi_t^{ij} dt + \sigma \left( |X_t^{(i,N)} - v_t^{(N)}| - |X_t^{(j,N)} - v_t^{(N)}| \right) dW_t^i + \sigma |X_t^{(j,N)} - v_t^{(N)}| dW_t^{ij},$$

where  $W_t^{ij} = W_t^i - W_t^j$  is again a standard Wiener process. Applying the Itô formula gives

$$\begin{aligned} \frac{d}{dt} \mathbb{E} |\xi_t^{ij}|^2 &\leq -2\lambda \mathbb{E} |\xi_t^{ij}|^2 + d\sigma^2 \mathbb{E} [(|X_t^{(i,N)} - v_t^{(N)}| - |X_t^{(j,N)} - v_t^{(N)}|)^2] + 2d\sigma^2 \mathbb{E} [|X_t^{(j,N)} - v_t^{(N)}|^2] \\ &\quad + d\sigma^2 \mathbb{E} [(|X_t^{(i,N)} - v_t^{(N)}| - |X_t^{(j,N)} - v_t^{(N)}|) |X_t^{(j,N)} - v_t^{(N)}|] \\ &\leq -\left(2\lambda - \frac{3}{2}d\sigma^2\right) \mathbb{E} |\xi_t^{ij}|^2 + \frac{5d\sigma^2}{2} \mathbb{E} [|X_t^{(j,N)} - v_t^{(N)}|^2], \end{aligned}$$

where we used the fact that

$$\left| |X_t^{(i,N)} - v_t^{(N)}| - |X_t^{(j,N)} - v_t^{(N)}| \right| \leq |X_t^{(i,N)} - X_t^{(j,N)}| = |\xi_t^{ij}|.$$

In order to estimate the last term on the right hand side, we first observe that

$$|X_t^{(j,N)} - v_t^{(N)}| \leq \frac{\sum_{\ell} \omega_f^{\alpha}(X_t^{(\ell,N)}) |X_t^{(j,N)} - X_t^{(\ell,N)}|}{\sum_{\ell} \omega_f^{\alpha}(X_t^{(\ell,N)})} \leq \sum_{\ell} \beta_t^{(\ell,N)} |\xi_t^{\ell j}| = \langle \beta_t^N, \zeta_t^j \rangle,$$

where

$$\beta_t^N = (\beta_t^{(1,N)}, \dots, \beta_t^{(N,N)})^{\top}, \quad \beta_t^{(\ell,N)} = \omega_f^{\alpha}(X_t^{(\ell,N)}) / \sum_{\ell} \omega_f^{\alpha}(X_t^{(\ell,N)}),$$

and  $\zeta^j = (|\xi^{1j}|, \dots, |\xi^{Nj}|)^{\top}$ . We further note that

$$\beta_t^N \geq 0, \quad \sum_{\ell} \beta_t^{(\ell,N)} = 1 = \sum_{\ell} e^{(\ell,N)}, \quad e^N = \left( \frac{1}{N}, \dots, \frac{1}{N} \right)^{\top}.$$

Moreover, majorization theory tells us that  $e^N \prec \beta_t^N$ , i.e.,  $\beta_t^N$  strongly majorizes  $e^N$  [30]. In this case, Horn's lemma provides the existence of an orthostochastic matrix  $A = (Q)^2$ ,  $Q$  orthogonal, such that  $e^N = A\beta_t^N$  [23]. Consequently,

$$\langle \beta_t^N, \zeta_t^j \rangle = \langle Q_t^{\top} e^N, Q_t \zeta_t^j \rangle \leq |Q_t^{\top} e^N|_2 |Q_t \zeta_t^j|_2 = |e^N|_2 |\zeta_t^j|_2 = \frac{1}{\sqrt{N}} \left( \sum_{\ell} |\xi_t^{\ell j}|^2 \right)^{\frac{1}{2}},$$

and therefore

$$(7) \quad |X_t^{(j,N)} - v_t^{(N)}|^2 \leq \frac{1}{N} \sum_{\ell} |\xi_t^{\ell j}|^2.$$

Finally, we obtain

$$(8) \quad \frac{d}{dt} \mathbb{E} |\xi_t^{ij}|^2 \leq -\left(2\lambda - \frac{3}{2}d\sigma^2\right) \mathbb{E} |\xi_t^{ij}|^2 + \frac{5d\sigma^2}{2N} \sum_{\ell} \mathbb{E} |\xi_t^{\ell j}|^2.$$

We now introduce the functions

$$S_t^{(N)} := \frac{1}{N} \sum_{i,j} \mathbb{E} |\xi_t^{ij}|^2, \quad R_t^{(N)} := \max_{1 \leq i,j \leq N} \mathbb{E} |\xi_t^{ij}|^2.$$

Then it follows from (8) that

$$(9) \quad S_t^{(N)} \leq \exp(-2\gamma t) S_0^{(N)}, \quad R_t^{(N)} \leq \exp(-2\gamma t) R_0^{(N)},$$

with  $\gamma = \lambda - 2d\sigma^2$ . From these observations, we have the following theorem.

**Theorem 2.2.** *Let  $\{\mathbf{X}_t^{(N)} \mid t > 0\}$  be a strong solution to the stochastic differential equation (5) with the initial data  $\mathbf{X}_0^{(N)}$  satisfying  $\mathbb{E} |\mathbf{X}_0^{(N)}|^2 < \infty$ . Then we have*

$$\mathbb{E} |\mathbf{X}_t^{(N)}|^2 \leq e^{\lambda t} \mathbb{E} |\mathbf{X}_0^{(N)}|^2 + \frac{\lambda + d\sigma^2}{\lambda + 2\gamma} S_0^{(N)} (e^{\lambda t} - e^{-2\gamma t}),$$

with  $\gamma = \lambda - 2d\sigma^2$ . Furthermore, if  $\gamma > 0$ , then we have

$$\max_{1 \leq i \leq N} \mathbb{E} [|X_t^{(i,N)} - v_t^{(N)}|^2] \rightarrow 0 \quad \text{as } t \rightarrow \infty,$$

exponentially fast with decay rate  $\gamma > 0$ .

*Proof.* Similarly as before, we get

$$\begin{aligned} \frac{d}{dt} \mathbb{E}|\mathbf{X}_t^{(N)}|^2 &= -2\lambda \mathbb{E}[\mathbf{X}_t^{(N)} \cdot \mathbf{F}_N(\mathbf{X}_t^{(N)})] + d\sigma^2 \mathbb{E}[|\mathbf{F}_N(\mathbf{X}_t^{(N)})|^2] \\ &\leq \lambda \mathbb{E}|\mathbf{X}_t^{(N)}|^2 + (\lambda + d\sigma^2) \mathbb{E}[|\mathbf{F}_N(\mathbf{X}_t^{(N)})|^2]. \end{aligned}$$

On the other hand we have from (9) that

$$\mathbb{E}[|\mathbf{F}_N(\mathbf{X}_t^{(N)})|^2] \leq \sum_i \mathbb{E}[|X_t^{(i,N)} - v_t^{(N)}|^2] \leq \frac{1}{N} \sum_{i,\ell} \mathbb{E}|\xi_t^{\ell i}|^2 \leq S_t^{(N)} \leq S_0^{(N)} e^{-2\gamma t}.$$

This yields

$$\mathbb{E}|\mathbf{X}_t^{(N)}|^2 \leq e^{\lambda t} \mathbb{E}|\mathbf{X}_0^{(N)}|^2 + \frac{\lambda + d\sigma^2}{\lambda + 2\gamma} S_0^{(N)} (e^{\lambda t} - e^{-2\gamma t}).$$

Now let  $\gamma > 0$ . Notice that (7) gives

$$\mathbb{E}[|X_t^{(i,N)} - v_t^{(N)}|^2] \leq \frac{1}{N} \sum_j |\xi_t^{ij}|^2 \leq R_t^{(N)}.$$

Together with (9), this gives

$$\max_{1 \leq i \leq N} \mathbb{E}[|X_t^{(i,N)} - v_t^{(N)}|^2] \leq R_0^{(N)} e^{-2\gamma t},$$

as  $t \rightarrow \infty$ . This completes the proof.  $\square$

*Remark 2.2.* Since  $\ell^p$ -norms are equivalent in finite dimensions for any  $1 \leq p \leq \infty$ , the initial condition  $\mathbb{E}|\mathbf{X}_0^{(N)}|^2 < \infty$  implies  $S_0^{(N)} \leq 4\mathbb{E}|\mathbf{X}_0^{(N)}|^2$  and  $R_0^{(N)} \leq 4\mathbb{E}|\mathbf{X}_0^{(N)}|^2$ .

### 3. WELL-POSEDNESS OF THE MEAN-FIELD EQUATION

In this section, we provide the well-posedness of the nonlocal, nonlinear Fokker–Planck equation (4). Since we will be working primarily with Borel probability measures on  $\mathbb{R}^d$  with finite second moment, we provide its definition for the readers convenience. We denote by

$$\mathcal{P}_2(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \quad \text{such that} \quad \int_{\mathbb{R}^d} |z|^2 \mu(dz) < \infty \right\}$$

to be the space of Borel probability measures on  $\mathbb{R}^d$  with finite second moment. This space may be equipped with the 2-Wasserstein distance  $W_2$  defined by

$$W_2^2(\mu, \hat{\mu}) = \inf_{\pi \in \Pi(\mu, \hat{\mu})} \left[ \int_{\mathbb{R}^d \times \mathbb{R}^d} |z - \hat{z}|^2 \pi(dz, d\hat{z}) \right], \quad \mu, \hat{\mu} \in \mathcal{P}_2(\mathbb{R}^d),$$

where  $\Pi(\mu, \hat{\mu})$  denotes the collection of all Borel probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\mu$  and  $\hat{\mu}$  on the first and second factors respectively. The set  $\Pi(\mu, \hat{\mu})$  is also known as the set of all couplings of  $\mu$  and  $\hat{\mu}$ . Equivalently, the Wasserstein distance may be defined by

$$W_2^2(\mu, \hat{\mu}) = \inf \mathbb{E} \left[ |Z - \hat{Z}|^2 \right],$$

where the infimum is taken over all joint distributions of the random variables  $Z$  and  $\hat{Z}$  with marginals  $\mu$  and  $\hat{\mu}$  respectively. It is known that  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$  is Polish, where  $W_2$  metricizes the weak convergence in  $\mathcal{P}_2(\mathbb{R}^d)$ , as well as, provides convergence of the first two moments [2, 36].

The main result of this section is provided by the following theorem.

**Theorem 3.1.** *Let  $f \in Lip_{loc}(\mathbb{R}^d)$ ,  $f \geq 0$  and  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Then there exists a unique nonlinear process  $\bar{X} \in \mathcal{C}([0, T], \mathbb{R}^d)$  satisfying (2) in the strong sense, and its law  $\rho_t = \text{law}(\bar{X}_t)$  satisfies the Fokker–Planck equation (4) with  $\rho \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$ ,  $\lim_{t \rightarrow 0} \rho_t = \rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ .*

*Proof.* For some given  $u \in \mathcal{C}([0, T], \mathbb{R}^d)$ , we may uniquely solve the stochastic differential equation

$$(10) \quad dX_t = -\lambda(X_t - u_t)dt + \sigma|X_t - u_t|dW_t, \quad \text{law}(X_0) = \rho_0,$$

for some fixed initial measure  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ , which induces  $\rho_t = \text{law}(X_t)$ . Since  $X \in \mathcal{C}([0, T], \mathbb{R}^d)$ , we obtain  $\rho \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$ , which satisfies the following Fokker–Planck equation

$$\frac{d}{dt} \int \varphi d\rho_t = \int \left( (\sigma^2/2)|x - u_t|^2 \Delta \varphi - \lambda(x - u_t) \cdot \nabla \varphi \right) d\rho_t \quad \text{for all } \varphi \in \mathcal{C}_b^2(\mathbb{R}^d).$$

We then compute  $v_f[\rho] \in \mathcal{C}([0, T], \mathbb{R}^d)$  according to (2b). This provides the self-mapping property of the map  $\mathcal{T}: u \mapsto v_f$ , which we now show to be a contraction in the metric space  $\mathcal{C}([0, T], \mathbb{R}^d)$ , endowed with the equivalent weighted norm

$$\|u\|_{\text{exp}} = \sup_{t \in [0, T]} |u_t| e^{\beta t},$$

for some  $\beta \in \mathbb{R}$  to be chosen appropriately later.

We begin by establishing an estimate of the second moment of  $\rho_t = \text{law}(\bar{X}_t)$  satisfying (10):

$$\begin{aligned} \frac{d}{dt} \int |x|^2 d\rho_t &= \int \left( d\sigma^2 |x - u_t|^2 - 2\lambda \langle x - u_t, x \rangle \right) d\rho_t \\ &= \int \left( (d\sigma^2 - 2\lambda)|x|^2 + (2\lambda - 2d\sigma^2) \langle x, u_t \rangle + d\sigma^2 |u_t|^2 \right) d\rho_t \\ &\leq -(2\lambda - d\sigma^2 - |\gamma|) \int |x|^2 d\rho_t + (d\sigma^2 + |\gamma|) |u_t|^2, \end{aligned}$$

with  $\gamma := \lambda - d\sigma^2$ . From Gronwall's inequality we deduce

$$(11) \quad \int |x|^2 d\rho_t \leq \left( \int |x|^2 d\rho_0 \right) e^{-(2\lambda - d\sigma^2 - |\gamma|)t} + \frac{d\sigma^2 + |\gamma|}{2\lambda - d\sigma^2 - |\gamma|} \left( 1 - e^{-(2\lambda - d\sigma^2 - |\gamma|)t} \right) \|u\|_{\infty}^2.$$

Consequently, we have  $\sup_{t \in [0, T]} \int |x|^2 d\rho_t \leq M$  for some  $M > 0$ , depending only on the coefficients  $\lambda, \sigma$ , the second moment of  $\rho_0$  and  $\|u\|_{\infty}$ . We then define the stopping time  $\tau_k := \inf\{t > 0 : |X_t| \geq k\}$  which reaches  $T$  for  $k$  large enough since the second moment is uniformly bounded in  $k$ .

Now let  $u, \hat{u} \in \mathcal{C}([0, T], \mathbb{R}^d)$  be given and  $X, \hat{X} \in \mathcal{C}([0, T], \mathbb{R}^d)$  their corresponding solutions to (10) respectively with  $\rho_t = \text{law}(X_t)$ ,  $\hat{\rho}_t = \text{law}(\hat{X}_t)$  and  $\text{law}(X_0) = \text{law}(\hat{X}_0) = \rho_0$ . Set  $\bar{\tau}_k := \tau_k \wedge \hat{\tau}_k$ . Taking the difference  $z_t := X_t - \hat{X}_t$  and applying the Itô formula on  $|z_t|^2$  gives

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[|z_t|^2] &= -2\lambda \mathbb{E}[|z_t|^2] + 2\lambda \mathbb{E}[\langle z_t, u_t - \hat{u}_t \rangle] + d\sigma^2 \mathbb{E}[(|X_t - u_t| - |\hat{X}_t - \hat{u}_t|)^2] \\ &\leq -(\lambda - 2d\sigma^2) \mathbb{E}[|z_t|^2] + (\lambda + 2d\sigma^2) |u_t - \hat{u}_t|^2. \end{aligned}$$

From the estimate above, we obtain via Gronwall's inequality

$$\begin{aligned} \mathbb{E}[|z_t|^2] &\leq \mathbb{E}[|z_0|^2] \leq (\lambda + 2d\sigma^2) e^{-(\lambda - 2d\sigma^2)t} \int_0^t e^{(\lambda - 2d\sigma^2)s} |u_s - \hat{u}_s|^2 ds \\ &\leq (\lambda + 2d\sigma^2) e^{-(\lambda - 2d\sigma^2)t} \int_0^t \left( e^{(\lambda - 2d\sigma^2 - \delta)s/2} |u_s - \hat{u}_s| \right)^2 e^{\delta s} ds \\ &\leq c_\delta e^{-2\beta t} \|u - \hat{u}\|_{\text{exp}}^2, \end{aligned}$$

where  $c_\delta := (\lambda + 2d\sigma^2)/\delta$  and we set  $\beta := (\lambda - 2d\sigma^2 - \delta)/2$ . This yields

$$(12) \quad \mathbb{E}[|z_t|] e^{\beta t} \leq \sqrt{c_\delta} \|u - \hat{u}\|_{\text{exp}}.$$

We next provide a stability estimate. Note that

$$v_f[\rho_t] = \frac{\int x \omega_f^\alpha d\rho_t}{\|\omega_f^\alpha\|_{L^1(\rho_t)}} = \frac{\mathbb{E}[\bar{X}_t \omega_f^\alpha(\bar{X}_t)]}{\mathbb{E}[\omega_f^\alpha(\bar{X}_t)]}.$$

Using the above equality together with taking care of the stopping time  $\bar{\tau}_k$  yields

$$|\omega_f^\alpha(x)| \leq 1, \quad \mathbb{E}[\omega_f^\alpha(\bar{X}_t)] \geq e^{-\alpha f_k}, \quad \text{and} \quad \mathbb{E}[|\omega_f^\alpha(\bar{X}_t) - \omega_f^\alpha(\hat{X}_t)|] \leq \alpha f_k \mathbb{E}[|\bar{X}_t - \hat{X}_t|],$$

for  $0 \leq t \leq \bar{\tau}_k$ , where  $f_k := \|f\|_{W^{1,\infty}(B_{2k})}$ . This allows us to obtain

$$\begin{aligned} |v_f[\rho_t] - v_f[\hat{\rho}_t]| &\leq \left| \frac{\mathbb{E}[\bar{X}_t(\omega_f^\alpha(\bar{X}_t) - \omega_f^\alpha(\hat{X}_t))]}{\mathbb{E}[\omega_f^\alpha(\bar{X}_t)]} \right| + \left| \frac{\mathbb{E}[(\bar{X}_t - \hat{X}_t)\omega_f^\alpha(\hat{X}_t)]}{\mathbb{E}[\omega_f^\alpha(\bar{X}_t)]} \right| \\ &\quad + \left| \frac{\mathbb{E}[\omega_f^\alpha(\hat{X}_t) - \omega_f^\alpha(\bar{X}_t)]}{\mathbb{E}[\omega_f^\alpha(\bar{X}_t)]\mathbb{E}[\omega_f^\alpha(\hat{X}_t)]} \mathbb{E}[\hat{X}_t\omega_f^\alpha(\hat{X}_t)] \right| \\ &\leq \alpha k f_k e^{\alpha f_k} \mathbb{E}[|\bar{X}_t - \hat{X}_t|] + e^{\alpha f_k} \mathbb{E}[|\bar{X}_t - \hat{X}_t|] + \alpha k f_k e^{2\alpha f_k} \mathbb{E}[|\bar{X}_t - \hat{X}_t|] \\ &\leq (e^{\alpha f_k} + 2\alpha k f_k e^{2\alpha f_k}) \mathbb{E}[|z_t|] \quad \text{for } 0 \leq t \leq \bar{\tau}_k. \end{aligned}$$

Since  $\bar{\tau}_k = T$  for  $k$  sufficiently large, we find that for some  $k_0$

$$|v_f[\rho_t] - v_f[\hat{\rho}_t]| \leq c_{k_0} \mathbb{E}[|z_t|] \quad \text{for } 0 \leq t \leq T,$$

where  $c_{k_0} := e^{\alpha f_{k_0}} + 2\alpha k_0 e^{2\alpha f_{k_0}} > 0$ . This together with (12) gives

$$\|v_f[\rho_t] - v_f[\hat{\rho}_t]\|_{\text{exp}} \leq c_{k_0} \sqrt{c_\delta} \|u - \hat{u}\|_{\text{exp}} \quad \text{for } 0 \leq t \leq T.$$

We finally choose  $\delta > 0$  large enough such that  $c_{k_0} \sqrt{c_\delta} < 1$  to have the contraction of the mapping. Applying the Banach fixed point theorem on the mapping  $\mathcal{T}$  concludes the proof.  $\square$

The following result appears as a simple consequence of Theorem 3.1.

**Corollary 3.1.** *Let  $\rho \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$  be the solution of the Fokker–Planck equation (4) with initial data  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$  provided by Theorem 3.1. Then  $v_f[\rho] \in \mathcal{C}([0, T], \mathbb{R}^d)$  and*

$$\sup_{t \in [0, T]} \int |x|^2 d\rho_t \leq \int |x|^2 d\rho_0 + \frac{d\sigma^2 + |\gamma|}{2\lambda - d\sigma^2 - |\gamma|} \|v_f[\rho]\|_{\infty}^2,$$

i.e., the second moment of  $\rho_t$  is uniformly bounded in the interval  $[0, T]$ .

#### 4. PROPAGATION OF CHAOS AND THE MEAN-FIELD LIMIT

Having the existence of a unique solution to the nonlinear process (2) at hand, we proceed to show a quantitative estimate (in  $N \in \mathbb{N}$ ) of the difference between the two solutions  $X_t^i$  and  $\bar{X}_t^i$  for any  $1 \leq i \leq N$ , and consequently also the difference between their respective laws. Unfortunately, the result holds under an additional assumption on the solution of the nonlinear process.

**Assumption 2.** The solution  $\rho = \text{law}(\bar{X}) \in \mathcal{C}([0, T], \mathcal{P}(\mathbb{R}^d))$  of (4) satisfies  $\|\omega_f^\alpha\|_{L^1(\rho_t)} \geq c$  for all  $t \in [0, T]$  for some constant  $c > 0$ .

*Remark 4.1.* In the next section, we provide a sufficient condition on  $f \in \text{Lip}_{\text{loc}}(\mathbb{R}^d)$  that will imply  $\|\omega_f^\alpha\|_{L^1(\rho_t)} \geq \|\omega_f^\alpha\|_{L^1(\rho_0)}$ , and therefore Assumption 2 (cf. Proposition 5.1).

**Proposition 4.1.** *Let  $f \in \text{Lip}_{\text{loc}}(\mathbb{R}^d)$ ,  $f \geq 0$ , and  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Consider the solution  $\mathbf{X}_t^{(N)}$  of (1) and the solution  $\bar{X}_t^{(i, N)}$ ,  $1 \leq i \leq N \in \mathbb{N}$  of (2) for  $t \in [0, T]$  satisfying Assumption 2, with mutually independent  $\rho_0$ -distributed initial data  $X_0^{(i, N)} = \bar{X}_0^{(i, N)}$ ,  $1 \leq i \leq N$ . Then the estimate*

$$(13) \quad \sup_{t \in [0, T]} \mathbb{E}[|X_t^{(i, N)} - \bar{X}_t^{(i, N)}|^2] \leq CN^{-1}, \quad 1 \leq i \leq N,$$

holds for some constant  $C > 0$ , independent of  $N \in \mathbb{N}$ .

*Proof.* Without loss of generality, we may take the same Wiener processes for both  $X_t^{(i, N)}$  and  $\bar{X}_t^{(i, N)}$  and define the sequence  $\{\tau_k\}_{k \geq 0}$  of stopping times

$$\tau_k := \inf\{t > 0 \mid |\bar{X}_t^{(1, N)}| \geq k\} \wedge \inf\{t > 0 \mid |\mathbf{X}_t^{(N)}| \geq k\}, \quad k \geq 0.$$

Since the second moment for both systems are uniformly bounded in  $k$  and  $N$ , we obtain  $\tau_k \geq T$  for  $k$  sufficiently large. Denoting the difference of the solutions by  $Z_t^{(i, N)} := X_{t \wedge \tau_k}^{(i, N)} - \bar{X}_{t \wedge \tau_k}^{(i, N)}$ , we



apply Itô's formula to obtain

$$\begin{aligned} \mathbb{E}[|Z_t^{(i,N)}|^2] &\leq 4\lambda^2 \mathbb{E} \left[ \left( \int_0^{t \wedge \tau_k} |X_s^{(i,N)} - \bar{X}_s^{(i,N)}| ds \right)^2 + \left( \int_0^{t \wedge \tau_k} |v_s^{(N)} - v_f[\rho_s]| ds \right)^2 \right] \\ &\quad + 4\sigma^2 \mathbb{E} \left[ \int_0^{t \wedge \tau_k} |X_s^{(i,N)} - \bar{X}_s^{(i,N)}|^2 ds + \int_0^{t \wedge \tau_k} |v_s^{(N)} - v_f[\rho_s]|^2 ds \right] \\ &\leq 4(\lambda^2 T + \sigma^2) \left( \int_0^{t \wedge \tau_k} \mathbb{E}[|Z_s^{(i,N)}|^2] ds + \int_0^{t \wedge \tau_k} \mathbb{E}[|v_s^{(N)} - v_f[\rho_s]|^2] ds \right). \end{aligned}$$

We remind the reader that

$$v_s^{(N)} = \left( \sum_i X_s^{(i,N)} \omega_f^\alpha(X_s^{(i,N)}) \right) / \left( \sum_i \omega_f^\alpha(X_s^{(i,N)}) \right).$$

In order to estimate the last term on the right-hand side, we write

$$v_s^{(N)} - v_f[\rho_s] = \frac{A + B - C + D}{N \|\omega_f^\alpha\|_{L^1(\rho_s)}},$$

and estimate each term on the right-hand side separately, where

$$\begin{aligned} A &= \sum_j (X_s^{(j,N)} - \bar{X}_s^{(j,N)}) \omega_f^\alpha(\bar{X}_s^{(j,N)}), \\ B &= \sum_j (X_s^{(j,N)} - v_s^{(N)}) (\omega_f^\alpha(X_s^{(j,N)}) - \omega_f^\alpha(\bar{X}_s^{(j,N)})), \\ C &= v_s^{(N)} \sum_j (\omega_f^\alpha(\bar{X}_s^{(j,N)}) - \mathbb{E}[\omega_f^\alpha(\bar{X}_s^{(j,N)})]), \\ D &= \sum_j (\bar{X}_s^{(j,N)} \omega_f^\alpha(\bar{X}_s^{(j,N)}) - \mathbb{E}[\bar{X}_s^{(j,N)} \omega_f^\alpha(\bar{X}_s^{(j,N)})]). \end{aligned}$$

We begin by computing the second moment of  $A$ . Simple computations lead to

$$\mathbb{E}[|A|^2] \leq N \sum_j \mathbb{E}[|Z_s^{(j,N)}|^2],$$

where we used Jensen's inequality. Note that for  $s \in (0, t \wedge \tau_k)$ , we have that  $|\bar{X}_s^{(1,N)}| \vee |\mathbf{X}_s^{(N)}| \leq k$  for some  $k \geq 0$ . Therefore, the estimate for  $B$  reads

$$\begin{aligned} \mathbb{E}[|B|^2] &\leq c_\omega(k)^2 \mathbb{E} \left[ \left( \sum_j |X_s^{(j,N)} - v_s^{(N)}|^2 \right) \left( \sum_j |Z_s^{(j,N)}|^2 \right) \right] \\ &\leq 2(1 + N)k^2 c_\omega(k)^2 \sum_j \mathbb{E}[|Z_s^{(j,N)}|^2], \end{aligned}$$

where  $c_\omega(k) = \|\omega_f^\alpha\|_{W^{1,\infty}(B_k)}$ . Now define  $\kappa_1(\bar{X}_s^{(i,N)}) = \omega_f^\alpha(\bar{X}_s^{(i,N)}) - \mathbb{E}[\omega_f^\alpha(\bar{X}_s^{(i,N)})]$ . Obviously  $\mathbb{E}[\kappa_1(\bar{X}_s^{(i,N)})] = 0$  for all  $1 \leq i \leq N$ . Moreover,

$$\mathbb{E}[\kappa_1(\bar{X}_s^{(j,N)}) \kappa_1(\bar{X}_s^{(k,N)})] = \mathbb{E}[\kappa_1(\bar{X}_s^{(j,N)})] \mathbb{E}[\kappa_1(\bar{X}_s^{(k,N)})] = 0 \quad \text{for } j \neq k,$$

since the processes  $\bar{X}_s^{(j,N)}$  and  $\bar{X}_s^{(k,N)}$  are independent for  $j \neq k$ . Consequently,

$$\mathbb{E}[|C|^2] \leq k^2 \sum_{j,k} \mathbb{E}[\kappa_1(\bar{X}_s^{(j,N)}) \kappa_1(\bar{X}_s^{(k,N)})] \leq N k^2 \mathbb{E}[\kappa_1(\bar{X}_s^{(1,N)})^2]$$

Similarly, we define  $\kappa_2(\bar{X}_s^{(i,N)}) = \bar{X}_s^{(i,N)} \omega_f^\alpha(\bar{X}_s^{(i,N)}) - \mathbb{E}[\bar{X}_s^{(i,N)} \omega_f^\alpha(\bar{X}_s^{(i,N)})]$ . As for  $\kappa_1$  we have that  $\mathbb{E}[\kappa_2(\bar{X}_s^{(i,N)})] = 0$  for all  $1 \leq i \leq N$  and  $\mathbb{E}[\kappa_2(\bar{X}_s^{(j,N)}) \kappa_2(\bar{X}_s^{(k,N)})] = 0$  for  $j \neq k$ , which yields

$$\mathbb{E}[|D|^2] \leq \sum_{j,k} \mathbb{E}[\kappa_2(\bar{X}_s^{(j,N)}) \kappa_2(\bar{X}_s^{(k,N)})] = N \mathbb{E}[\kappa_2(\bar{X}_s^{(1,N)})^2].$$

Note that  $\mathbb{E}[\kappa_1(\bar{X}_t^{(1,N)})^2]$  and  $\mathbb{E}[\kappa_2(\bar{X}_t^{(1,N)})^2]$  are uniformly bounded in time. Indeed, we find

$$\begin{aligned}\mathbb{E}[\kappa_1(\bar{X}_s^{(1,N)})^2] &= \int |x \omega_f^\alpha(x) - \mathbb{E}[\bar{X}_s^{(1,N)} \omega_f^\alpha(\bar{X}_s^{(1,N)})]|^2 d\rho_s \\ &= \int |x \omega_f^\alpha(x)|^2 d\rho_s - |\mathbb{E}[\bar{X}_s^{(1,N)} \omega_f^\alpha(\bar{X}_s^{(1,N)})]|^2 \\ &\leq \int |x \omega_f^\alpha(x)|^2 d\rho_s(x) \leq \int |x|^2 d\rho_s(x).\end{aligned}$$

Therefore, the uniform boundedness (in  $s$ ) follows from Corollary 3.1. We obtain a similar estimate for  $\kappa_2$ . By assumption  $\|\omega_f^\alpha\|_{L^1(\rho_s)} \geq c$  for all  $s \in [0, T]$ . Hence, we obtain

$$(14) \quad \mathbb{E}[|z_t^{(i,N)}|^2] \leq c_1 \int_0^{t \wedge \tau_k} \mathbb{E}[|z_s^{(i,N)}|^2] ds + c_2 k^2 \frac{1}{N} \sum_j \int_0^{t \wedge \tau_k} \mathbb{E}[|z_s^{(j,N)}|^2] ds + c_3 \frac{1}{N} t$$

with constants  $c_i \geq 0$ . Averaging over  $1 \leq i \leq N$  gives

$$y_t^{(N)} := \frac{1}{N} \sum_i \mathbb{E}[|z_t^{(i,N)}|^2] \leq c_0(1+k^2) \int_0^{t \wedge \tau_k} y_s^{(N)} ds + c_0 \frac{1}{N} t,$$

for some nonnegative constant  $c_0 \geq 0$ . An application of the Gronwall inequality yields

$$y_t^{(N)} \leq (c_0/N) t e^{c_0(1+k^2)t} \quad \text{for all } t \in [0, T].$$

Finally, we insert this estimate into (14) and use Gronwall's inequality again to obtain

$$\mathbb{E}[|X_{t \wedge \tau_k}^{(i,N)} - \bar{X}_{t \wedge \tau_k}^{(i,N)}|^2] = \mathbb{E}[|z_t^{(i,N)}|^2] \leq \frac{c_0}{(1+k^2)} \frac{1}{N} T \left( k^2 e^{c_0(1+k^2)T} + 1 \right) e^{c_0 T}.$$

Since the estimate above holds for any  $k \in \mathbb{N}$ , and  $\tau_{k_0} \geq T$  for some  $k_0 \in \mathbb{N}$ , we finally obtain

$$\mathbb{E}[|X_t^{(i,N)} - \bar{X}_t^{(i,N)}|^2] \leq \frac{c_0}{(1+k_0^2)} \frac{1}{N} T \left( k_0^2 e^{c_0(1+k_0^2)T} + 1 \right) e^{c_0 T} =: CN^{-1},$$

which is precisely the required estimate for any  $1 \leq i \leq N$ .  $\square$

*Remark 4.2.* The moment estimate (13) ensures (i) the so-called *propagation of chaos* property [33], as well as (ii) the convergence of the stochastic empirical measure

$$\rho_t^N(A) = \frac{1}{N} \sum_i \delta_{X_t^i}(A), \quad A \in \mathcal{B}(\mathbb{R}^d),$$

towards the deterministic mean-field probability measure  $\rho_t \in \mathcal{P}_2(\mathbb{R}^d)$  [35].

Indeed, for (i) we have that

$$W_2^2(\rho_t^{(\ell,N)}, \rho_t^{\otimes \ell}) \leq \mathbb{E}[|(X_t^1, \dots, X_t^\ell) - (\bar{X}_t^1, \dots, \bar{X}_t^\ell)|^2] \leq \ell CN^{-1}, \quad 1 \leq \ell \leq N,$$

where  $\rho_t^{(\ell,N)} \in \mathcal{P}_2(\mathbb{R}^{\ell d})$  denotes the  $\ell$ -th marginal of the  $N$ -particle joint distribution on  $\mathbb{R}^{Nd}$ .

As for (ii), we have for any  $\varphi \in \text{Lip}_b(\mathbb{R}^d)$  the estimate

$$\begin{aligned}\mathbb{E} \left[ \left| \frac{1}{N} \sum_i \varphi(X_t^i) - \int \varphi d\rho_t \right|^2 \right] \\ \leq 2\mathbb{E} \left[ \frac{1}{N} \sum_i |\varphi(X_t^i) - \varphi(\bar{X}_t^i)|^2 + \left| \frac{1}{N} \sum_i \varphi(\bar{X}_t^i) - \int \varphi d\rho_t \right|^2 \right] \leq \tilde{C} N^{-1},\end{aligned}$$

for some constant  $\tilde{C}$ , independent of  $N$  and  $t \in [0, T]$ . Note that the estimate for second term in the first inequality follows from the law of large numbers, which holds since  $\{\bar{X}_t^i\}$  are mutually independent and identically  $\rho_t$ -distributed.

Summarizing the above discussion, we have the following result.

**Theorem 4.1.** *Let  $\rho_t^{(\ell, N)}$ ,  $\ell \leq N$  be the  $\ell$ -th marginal of the  $N$ -particle joint distribution given by  $\text{law}(X_t^1, \dots, X_t^N)$  on  $\mathbb{R}^{Nd}$  and  $\rho_t = \text{law}(\bar{X}_t^i)$ , where  $(X_t^1, \dots, X_t^N)$  solves (1) and  $\bar{X}_t^i$  solves (2) with initial data  $\text{law}(X_0^i) = \text{law}(\bar{X}_0^i) = \rho_0$ . Then for some arbitrary, but fixed  $\ell \geq 1$ :*

$$\sup_{t \in [0, T]} W_2(\rho_t^{(\ell, N)}, \rho_t^{\otimes \ell}) \longrightarrow 0 \quad \text{for } N \rightarrow \infty.$$

*In particular, the first marginal  $\rho_t^{(1, N)} \in \mathcal{P}_2(\mathbb{R}^d)$  converges in law towards the 1-particle distribution  $\rho_t \in \mathcal{P}_2(\mathbb{R}^d)$  uniformly in time  $t \in [0, T]$  as  $N \rightarrow \infty$ . Furthermore, the Wasserstein distance  $W_2$  ensures that their first and second moment coincide in the limit.*

## 5. LARGE TIME BEHAVIOR AND CONSENSUS FORMATION

We finally arrive the most important part of the paper, i.e., we provide sufficient conditions such that *uniform consensus formation* happens. More precisely, we say that uniform consensus formation occurs when

$$\rho_t \longrightarrow \delta_{\hat{x}} \quad \text{as } t \rightarrow \infty,$$

for some  $\hat{x} \in \mathbb{R}^d$  possibly depending on  $\rho_0$ . In fact, in the framework of global optimization, we would like to further have that  $\hat{x} = x_* = \inf f$ . In other words, we want that  $\rho_t$  concentrates at the global minimum of  $f$ . Unlike, the deterministic case, the formation of non-uniform consensus, i.e., multiple opinions in the limit  $t \rightarrow \infty$ , in the stochastic model cannot occur [31]. Hence, it is expected that uniform consensus is formed, whenever concentration happens. We will see that this is the case.

**5.1. Concentration estimates.** Let us assume for the moment that  $\rho \in \mathcal{C}([0, \infty), \mathcal{P}_2(\mathbb{R}^d))$  solves the Fokker–Planck equation (4). Then, we may estimate its expectation explicitly via

$$(15) \quad \frac{d}{dt} E(\rho_t) := \frac{d}{dt} \int x d\rho_t = -\lambda \int (x - v_f[\rho_t]) d\rho_t.$$

Furthermore, simple computation of the evolution of its variance gives

$$(16) \quad \frac{d}{dt} V(\rho_t) := \frac{d}{dt} \frac{1}{2} \int |x - E(\rho_t)|^2 d\rho_t = -2\lambda V(\rho_t) + (d\sigma^2/2) \int |x - v_f[\rho_t]|^2 d\rho_t,$$

To estimate the last term on the right, we apply Jensen's inequality again to obtain

$$(17) \quad \int |x - v_f[\rho_t]|^2 d\rho_t \leq \frac{\iint |x - y|^2 \omega_f^\alpha d\rho_t(x) d\rho_t(y)}{\int \omega_f^\alpha d\rho_t} \leq 2e^{-\alpha f_*} V(\rho_t) / \|\omega_f^\alpha\|_{L^1(\rho_t)},$$

where we used the assumption  $f_* = \inf f \geq 0$ .

If  $f \in \text{Lip}_b(\mathbb{R}^d)$ , i.e.,  $f$  is additionally bounded, then a rough estimate would be to use the fact that  $\|\omega_f^\alpha\|_{L^1(\rho_t)} \geq e^{-\alpha \sup f}$ . In this case we have

$$\frac{d}{dt} V(\rho_t) \leq -(2\lambda - d\sigma^2 \text{osc}(\omega_f^\alpha)) V(\rho_t),$$

which guarantees concentration when  $2\lambda > d\sigma^2 \text{osc}(\omega_f^\alpha)$ , due to Gronwall's inequality. This estimate is rather poor since  $\text{osc}(\omega_f^\alpha) \rightarrow \infty$  as  $\alpha \rightarrow \infty$ . Fortunately, we can do better.

We begin by providing an estimate for  $\|\omega_f^\alpha\|_{L^1(\rho_t)}$  when imposing further regularity of  $f$ . To do so, we make use of the following equality

$$(18) \quad \begin{aligned} \frac{d}{dt} \log \left( \int e^{-\alpha f} d\rho_t \right) &= \alpha \lambda \int \langle x - v_f[\rho_t], \nabla f(x) - \nabla f(v_f[\rho_t]) \rangle d\eta_t^\alpha \\ &\quad - (\alpha \sigma^2 / 2) \int (\Delta f - \alpha |\nabla f|^2) |x - v_f[\rho_t]|^2 d\eta_t^\alpha, \end{aligned}$$

where we used the fact that

$$\int (x - v_f[\rho_t]) d\eta_t^\alpha = 0,$$

for the first term. At this point, we can provide a sufficient condition for concentration to happen. Indeed, assuming that the right-hand side of (18) to be nonnegative, we deduce that

$$\|\omega_f^\alpha\|_{L^1(\rho_t)} = \int e^{-\alpha f} d\rho_t \geq \left( \int e^{-\alpha f} d\rho_0 \right) = \|\omega_f^\alpha\|_{L^1(\rho_0)},$$

Consequently, we have that

$$\frac{d}{dt} V(\rho_t) \leq - \left( 2\lambda - d\sigma^2 (e^{-\alpha f^*} / \|\omega_f^\alpha\|_{L^1(\rho_0)}) \right) V(\rho_t).$$

Notice that for any  $\rho_0 \in \mathcal{P}_2^{ac}(\mathbb{R}^d)$  containing  $x_*$  in its support, it follows from the *Laplace principle* (cf. [31]) that for any  $\varepsilon > 0$  there exists some  $\bar{\alpha} > 0$  such that

$$\|\omega_f^\alpha\|_{L^1(\rho_0)} \geq (1 + \varepsilon) e^{-\alpha f^*} \quad \text{for any } \alpha \geq \bar{\alpha},$$

and this implies

$$(19) \quad \frac{e^{-\alpha f^*}}{\|\omega_f^\alpha\|_{L^1(\rho_0)}} \leq (1 + \varepsilon) = \mathcal{O}(1).$$

Combining (16), (17) and (19) together with assumption that the right-hand side of (18) remains nonnegative for all times, we obtain the estimate

$$(20) \quad \frac{d}{dt} V(\rho_t) \leq - (2\lambda - (1 + \varepsilon)d\sigma^2) V(\rho_t).$$

Therefore, by choosing  $\lambda > 0$  sufficiently large, we have the desired concentration estimate.

The concentration estimate (20) provides the existence of some point  $\hat{x} \in \mathbb{R}^d$ , possibly depending on  $\rho_0$ , such that  $E(\rho_t) \rightarrow \hat{x}$  as  $t \rightarrow \infty$ . Indeed, we obtain from (17) that

$$|E(\rho_t) - v_f[\rho_t]|^2 \leq \int |x - v_f[\rho_t]|^2 d\rho_t \leq 2(1 + \varepsilon)V(\rho_t) \leq 2(1 + \varepsilon)V(\rho_0)e^{-2\lambda_\varepsilon t},$$

with  $\lambda_\varepsilon = \lambda - (1 + \varepsilon)d\sigma^2/2$ . On the other hand, we have from (15) the estimate

$$\begin{aligned} \int_0^t \left| \frac{d}{dt} E(\rho_t) \right| dt &\leq \lambda \int_0^t \int |x - v_f[\rho_s]| d\rho_s ds \leq 2\lambda(1 + \varepsilon)V(\rho_0) \int_0^t e^{-2\lambda_\varepsilon s} ds \\ &= (\lambda/\lambda_\varepsilon)(1 + \varepsilon)V(\rho_0)(1 - e^{-\lambda_\varepsilon t}), \end{aligned}$$

which tells us that  $\frac{d}{dt} E(\rho_t) \in L^1(0, \infty)$  and thus, there exists some point  $\hat{x} \in \mathbb{R}^d$ , possibly depending on  $\rho_0$ , such that

$$\hat{x} = E(\rho_0) + \int_0^\infty \frac{d}{dt} E(\rho_t) dt = \lim_{t \rightarrow \infty} E(\rho_t) \quad \text{with } |\hat{x} - E(\rho_0)| \leq (\lambda/\lambda_\varepsilon)(1 + \varepsilon)V(\rho_0).$$

Furthermore, since  $|E(\rho_t) - v_f[\rho_t]| \rightarrow 0$  for  $t \rightarrow \infty$ , we have that  $\lim_{t \rightarrow \infty} v_f[\rho_t] = \hat{x}$ .

We now provide a sufficient condition for the right-hand side of (18) to be nonnegative.

**Proposition 5.1.** *Let  $f \in Lip_{loc}(\mathbb{R}^d)$ ,  $f \geq 0$ , satisfy the following additional conditions:*

- (i)  *$f$  may be expressed as the sum of two functions*

$$f(x) = g(x) + h(x),$$

*where  $g \in \mathcal{C}^2(\mathbb{R}^d)$  is globally convex, i.e.,*

$$\langle \nabla g(x) - \nabla g(y), x - y \rangle \geq c_g |x - y|^2 \quad \forall x, y \in \mathbb{R}^d,$$

*for some constant  $c_g > 0$ , and  $h \in \mathcal{C}^2(\mathbb{R}^d)$  satisfies  $\|\nabla h\|_{Lip} =: c_h < c_g$ .*

- (ii) *There exist constants  $c_0, c_f > 0$ , such that*

$$\Delta f \leq c_0 + c_f |\nabla f|^2 \quad \text{in } \mathbb{R}^d.$$

*Then there exists  $\alpha$  and  $\lambda > 0$  such that the right-hand side of (18) remains nonnegative for all times  $t > 0$ . In particular, we obtain uniform consensus for  $\rho_t$  as  $t \rightarrow \infty$ .*

*Proof.* Under the assumptions on  $f$  above, we deduce

$$\begin{aligned}
 \frac{d}{dt} \log \left( \int e^{-\alpha f} d\rho_t \right) &= \alpha \lambda \int \langle x - v_f[\rho_t], \nabla g(x) - \nabla g(v_f[\rho_t]) \rangle d\eta_t^\alpha \\
 &\quad + \alpha \lambda \int \langle x - v_f[\rho_t], \nabla h(x) - \nabla h(v_f[\rho_t]) \rangle d\eta_t^\alpha \\
 &\quad - (\alpha \sigma^2 / 2) \int (\Delta f - \alpha |\nabla f|^2) |x - v_f[\rho_t]|^2 d\eta_t^\alpha \\
 &\geq \alpha \lambda (c_g - c_h) \int |x - v_f[\rho_t]|^2 d\eta_t^\alpha - (\alpha \sigma^2 / 2) c_0 \int |x - v_f[\rho_t]|^2 d\eta_t^\alpha \\
 &\quad + (\alpha \sigma^2 / 2) (\alpha - c_f) \int |\nabla f|^2 |x - v_f[\rho_t]|^2 d\eta_t^\alpha.
 \end{aligned}$$

Consequently, this yields

$$\begin{aligned}
 (21) \quad \frac{d}{dt} \log \left( \int e^{-\alpha f} d\rho_t \right) &\geq \alpha (\lambda (c_g - c_h) - (\sigma^2 / 2) c_0) \int |x - v_f[\rho_t]|^2 d\eta_t^\alpha \\
 &\quad + (\alpha \sigma^2 / 2) (\alpha - c_f) \int |\nabla f|^2 |x - v_f[\rho_t]|^2 d\eta_t^\alpha.
 \end{aligned}$$

Choosing  $\alpha \geq \alpha_0$  and  $\lambda \geq \lambda_0$ , where

$$\alpha_0 = c_f \quad \text{and} \quad \lambda_0 = \frac{1}{2} \frac{c_0 \sigma^2}{(c_g - c_h)},$$

provides the required nonnegativity of the right-hand side of (21).  $\square$

*Remark 5.1.* In the simple case  $f(x) = |x|^2/2$  clearly satisfies the requirements of Proposition 5.1 with  $\alpha_0 = 0$  and  $\lambda_0 = d\sigma^2/2$ , since  $c_g = 1$ ,  $c_h = c_f = 0$  and  $c_0 = d$ .

*Remark 5.2.* The assumptions in Proposition 5.1 on  $f$  also provide a sufficient condition to apply Proposition 4.1 without requiring Assumption 2.

**5.2. Approximate global minimizer.** While the previous results provided a sufficient condition for uniform consensus to occur, we will argue further that the point of consensus  $\hat{x} \in \mathbb{R}^d$  may be made arbitrarily close to the global minimum  $x_*$  of  $f$ , for  $f \in \text{Lip}_{loc}(\mathbb{R}^d)$  satisfying the assumptions in Proposition 5.1. Indeed, under the assumptions of Proposition 5.1, there exists some  $\alpha \gg 1$  and  $\lambda > 0$  such that the right-hand side of (18) remains nonnegative for all times  $t > 0$ . Therefore,

$$-\log \left( \int e^{-\alpha f} d\rho_t \right) \leq -\log \left( \int e^{-\alpha f} d\rho_0 \right) \quad \text{for all } t \geq 0.$$

Differentiating the inequality above w.r.t.  $\alpha$ , we obtain

$$\int f d\eta_t^\alpha \leq \int f d\eta_0^\alpha.$$

Now let  $\epsilon > 0$  be arbitrary but fixed. Due to the Laplace principle [31], we further obtain some  $\hat{\alpha} \gg 1$ , such that

$$\int (f - f_*) d\eta_t^{\hat{\alpha}} \leq \int (f - f_*) d\eta_0^{\hat{\alpha}} \leq \epsilon.$$

Then  $f(v_f[\rho_t])$  satisfies the following estimate:

$$\begin{aligned}
 f(v_f[\rho_t]) &= g(v_f[\rho_t]) + h(v_f[\rho_t]) \leq \int g(x) d\eta_t^{\hat{\alpha}} + h(v_f[\rho_t]) \\
 &= \int f(x) d\eta_t^{\hat{\alpha}} + \int (h(v_f[\rho_t]) - h(x)) d\eta_t^{\hat{\alpha}} \leq \int f(x) d\eta_t^{\hat{\alpha}} + \|\nabla f\|_{\sup} \int |x - v_f[\rho_t]| d\eta_t^{\hat{\alpha}},
 \end{aligned}$$

where we made use of Jensen's inequality in the first inequality. Consequently,

$$f(v_f[\rho_t]) - f_* \leq \epsilon + \|\nabla f\|_{\sup} \int |x - v_f[\rho_t]| d\eta_t^{\hat{\alpha}}.$$

Since the last term on the right-hand side converges to zero as  $t \rightarrow \infty$ , we may pass to the limit to obtain  $\lim_{t \rightarrow \infty} f(v_f[\rho_t]) \leq f_* + \epsilon$ . Due to the continuity of  $f$ , we find some  $\epsilon_0 > 0$  such that  $\lim_{t \rightarrow \infty} v_f[\rho_t] = \hat{x} \in B_{\epsilon_0}(x_*)$ , where  $\hat{x}$  should be chosen larger if necessary.

Summarizing the discussion above, we obtain the following proposition.

**Proposition 5.2.** *Let  $f \in \text{Lip}_{\text{loc}}(\mathbb{R}^d)$  satisfy the requirements of Proposition 5.1 and  $0 < \epsilon_0 \ll 1$  be arbitrarily small. Then there exists some  $\alpha \gg 1$  such that uniform consensus is obtained at a point  $\hat{x} \in B_{\epsilon_0}(x_*)$ . More precisely, we have that  $\rho_t \rightarrow \delta_{\hat{x}}$  for  $t \rightarrow \infty$ , with  $\hat{x} \in B_{\epsilon_0}(x_*)$ .*

## 6. 1D CASE: PSEUDO-INVERSE DISTRIBUTION AND EXTENDED MODELS

In this section, we consider the Fokker–Planck equation (4) in one spatial dimension and derive an equivalent formulation of the equation in terms of the pseudo-inverse distribution. Then, we introduce an extension of the current model to replace the diffusion term with nonlinear diffusions of porous media type, which would guarantee compact support of the probability measure  $\rho_t$ .

**6.1. Evolution of the inverse distribution function.** We first define the well-known cumulative distribution  $F_t$  of a probability measure  $\rho_t$  and its pseudo-inverse  $\chi_t$  (cf. [36]) by

$$F_t(x) = \int_{-\infty}^x d\rho_t = \rho_t((-\infty, x]).$$

Then, the pseudo-inverse of  $F_t$  on the interval  $[0, 1]$  is defined by

$$\chi_t(\eta) := F_t^{-1}(\eta) := \inf\{x \in \mathbb{R} \mid F_t(x) > \eta\}.$$

Both,  $F_t$  and  $\chi_t$  are by definition right-continuous. To derive the evolution equation for  $\chi_t$  we use the properties of  $\rho_t$ ,  $F_t$  and  $\chi_t$  collected in the following corollary. These properties can be obtained by basic calculus using the above definitions.

**Corollary 6.1.** *Let  $\rho_t$  be a probability measure,  $F_t$  the corresponding cumulative distribution and  $\chi$  the pseudo-inverse distribution of  $F$  as defined above. Then, the following equations hold*

$$\begin{aligned} \chi(t, F(t, x)) &= x, & F(t, \chi(t, \eta)) &= \eta, & \partial_t F &= -\rho \partial_t \chi, & \partial_\eta \chi &= \frac{1}{\rho}, \\ \partial_\eta \chi \partial_x F &= 1, & \frac{\partial_x \rho}{\rho} &= \frac{\partial_{xx} F}{\partial_x F} = -\frac{\partial_{\eta\eta} \chi}{(\partial_\eta \chi)^2} \end{aligned}$$

restricted to  $x = \chi(t, \eta)$ ,  $\eta = F(t, x)$  and  $x \in \text{supp}(\rho)$ , respectively.

From these properties we may now derive an integro-differential equation for the pseudo-inverse  $\chi_t$ ,  $t \geq 0$ . Indeed, let us consider the solution  $\rho_t$  to (4) satisfying  $\rho_t \in \mathcal{C}([0, \infty), \mathcal{P}_2(\mathbb{R}^d))$ . From the definition of  $F_t$ , we deduce that

$$\partial_t F_t(x) - \mu_t(x) \rho_t(x) = \rho_t(x) \partial_x \kappa_t(x) + \kappa_t(x) \partial_x \rho_t(x).$$

We then use the relations between  $\chi_t$  and  $\rho_t$  provided in Corollary 6.1 to obtain

$$-\rho_t(x) \partial_t \chi_t(\eta) - \mu_t(x) \rho_t(x) = \rho_t(x) \partial_x \kappa_t(x) + \kappa_t(x) \partial_x \rho_t(x),$$

which consequently yields

$$\partial_t \chi_t(\eta) + \mu_t(x) = -\partial_x \kappa_t(x) - \kappa_t(x) \frac{\partial_x \rho_t(x)}{\rho_t(x)} = -\frac{\partial_\eta \kappa_t(\eta)}{\partial_\eta \chi_t(\eta)} + \kappa_t(\eta) \frac{\partial_{\eta\eta} \chi_t(\eta)}{(\partial_\eta \chi_t(\eta))^2}.$$

On the other hand, it follows from Corollary 6.1 that  $\mu_t$  and  $\kappa_t$  may be rewritten as

$$\mu_t = \lambda(\chi_t - v_f[\chi]) \quad \text{and} \quad \kappa_t = (\sigma^2/2)|\chi_t - v_f[\chi]|^2,$$

respectively, where  $v_f[\chi]$  is given by

$$v_f[\chi] = \frac{\int_0^1 \chi_t \exp(-\alpha f(\chi_t(\eta))) d\eta}{\int_0^1 \exp(-\alpha f(\chi_t(\eta))) d\eta}.$$

Hence, the pseudo-inverse distribution  $\chi_t$  satisfies the following integro-differential equation:

$$(22) \quad \partial_t \chi_t + \mu_t = -\partial_\eta (\kappa_t (\partial_\eta \chi_t)^{-1}).$$

6.1.1. *Concentration estimates and consensus formation.* Using similar arguments as in Section 5, we obtain the large time behavior of solutions  $\chi_t$  to (22). For this, we set

$$V(\chi_t) := \frac{1}{4} \int_0^1 \int_0^1 |\chi_t(\eta) - \chi_t(\hat{\eta})|^2 d\eta d\hat{\eta}.$$

**Proposition 6.1.** *Let  $\chi_t$  satisfy equation (22). If  $\lambda > \sigma^2$ , then there exists  $\alpha_0 > 0$  such that*

$$V(\chi_t) \rightarrow 0 \quad \text{as } t \rightarrow \infty \quad \text{exponentially fast,}$$

*for all  $\alpha \geq \alpha_0$ . Furthermore, there exists some  $\hat{x} \in \mathbb{R}$  such that*

$$\chi_t(\eta) \rightarrow \hat{x} \quad \text{as } t \rightarrow \infty,$$

*for all  $\eta \in [0, 1]$ . In particular,  $\chi_t$  converges to the constant function  $\hat{x}$  in the limit  $t \rightarrow \infty$ .*

*Proof.* Using arguments similar to (16) and (17), we find that

$$\frac{d}{dt} V(\chi_t) = -2\lambda V(\chi_t) + (\sigma^2/2) \int_0^1 |\chi_t(\eta) - v_f(\chi_t)|^2 d\eta \leq -2 \left( \lambda - \frac{(\sigma^2/2)e^{-\alpha f_*}}{\int_0^1 e^{-\alpha f(\chi_t(\eta))} d\eta} \right) V(\chi_t),$$

where  $f_* = \inf f \geq 0$ . On the other hand, we get the following equality:

$$\begin{aligned} & \frac{d}{dt} \log \left( \int_0^1 e^{-\alpha f(\chi_t(\eta))} d\eta \right) \\ (23) \quad &= \alpha \lambda \int_0^1 (\chi_t(\eta) - v_f(\chi_t)) (f'(\chi_t(\eta)) - f'(v_f(\chi_t))) e^{-\alpha f(\chi_t(\eta))} d\eta \\ & \quad - \frac{\alpha \sigma^2}{2} \int_0^1 (f''(\chi_t(\eta)) - \alpha (f'(\chi_t(\eta)))^2) |\chi_t(\eta) - v_f(\chi_t)|^2 e^{-\alpha f(\chi_t(\eta))} d\eta. \end{aligned}$$

Assuming the right-hand side of the equality (23) to be positive, we obtain

$$\int_0^1 e^{-\alpha f(\chi_t(\eta))} d\eta \geq \int_0^1 e^{-\alpha f(\chi_0(\eta))} d\eta.$$

Notice that we can obtain the inequality above if  $f$  satisfies the conditions in Proposition 5.1. Finally, we employ again a similar strategy discussed in Section 5 to have that for any  $\epsilon > 0$  there exists a  $\alpha_0 > 0$  such that

$$\frac{d}{dt} V(\chi_t) \leq -2 \left( \lambda - \frac{(\sigma^2/2)e^{-\alpha f_*}}{\int_0^1 e^{-\alpha f(\chi_0(\eta))} d\eta} \right) V(\chi_t) \leq -2 (\lambda - (1 + \epsilon)\sigma^2) V(\chi_t) \quad \text{for } \alpha \geq \alpha_0.$$

This, together with the concentration estimates in subsection 5.1, completes the proof.  $\square$

**6.2. Porous media version of the evolution equation.** One very common application of the pseudo-inverse distribution  $\chi_t$  is to study the behavior of the support  $\text{supp}(\rho_t)$  of the corresponding probability measure  $\rho_t$ . This is especially interesting when  $\rho_t$  has compact support. Unfortunately, we do not have that in the present case due to the diffusion, which causes  $\rho_t$  to have full support in  $\mathbb{R}^d$ . This naturally leads to the idea of increasing the power of  $\rho_t$  in the diffusion term, inspired by the porous media equation [15]. The evolution equation for  $\rho_t$  then becomes

$$(24) \quad \partial_t \rho_t + \partial_x(\mu_t \rho_t) = \partial_{xx}(\kappa_t \rho_t^m).$$

with porous media coefficient  $m \geq 1$ . Notice that the previous model is included here for  $m = 1$ . The derivation of the evolution equation for  $\chi_t$  corresponding to this equation may be analogously done, which leads to

$$(25) \quad \partial_t \chi_t + \mu_t = -\partial_\eta(\kappa_t (\partial_\eta \chi_t)^{-m}).$$

Further investigation of the diffusion term results in

$$\partial_\eta(\kappa_t (\partial_\eta \chi_t)^{-m}) = \partial_\eta(\kappa_t \rho_t^m) = \rho_t^m \partial_\eta \kappa_t + \kappa_t \partial_\eta(\rho_t^m) = \rho_t^m \partial_\eta \kappa_t + m \kappa_t \rho_t^{m-1} \partial_\eta \rho_t$$

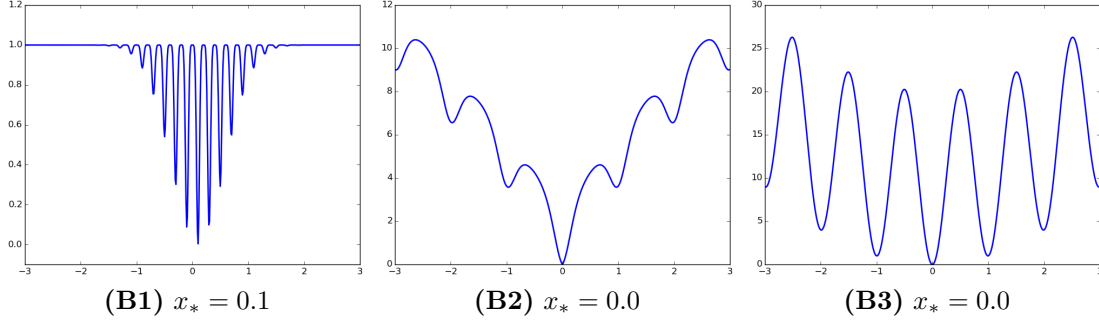


FIGURE 1. Benchmark functions for the numerical simulations.

in  $(\eta, t)$  variables. For  $m > 1$  we can do the following formal computations. Due to mass conservation of  $\rho_t$  we assume a no flux condition for (24) which in  $(x, t)$  variables reads

$$\mu_t \rho_t = \partial_x (\kappa_t \rho_t^m) = \rho_t^m \partial_x \kappa_t + m \kappa_t \rho_t^{m-1} \partial_x \rho_t,$$

on the boundary points of  $\text{supp}(\rho_t)$ . Consequently, we obtain

$$\partial_\eta (\kappa_t (\partial_\eta \chi_t)^{-m}) (F_t(x)) = [\rho_t^m \partial_x \kappa_t + m \kappa_t \rho_t^{m-1} \partial_x \rho_t] \partial_\eta \chi_t (F_t(x)) = \mu_t \rho_t \partial_\eta \chi_t (F_t(x)) = \mu_t,$$

on the boundary points of  $\text{supp}(\rho_t)$ . Therefore, restricting (25) onto the boundary points yields

$$(26) \quad \partial_t \chi_t(\eta) = -2\mu_t(\eta) = -2\lambda(\chi_t(\eta) - v_f[\chi]) \quad \text{for } \eta \in \{0, 1\}.$$

Since  $v_f[\chi]$  is always contained in the interior of  $\text{supp}(\rho_t)$  by definition,  $\mu_t$  is negative at the left boundary point  $\eta = 0$  and positive at the right boundary point  $\eta = 1$ . Hence, (26) implies the shrinking of  $\text{supp}(\rho_t)$  and in particular concentration of  $\chi_t$  at  $v_f$  as  $t \rightarrow \infty$ .

## 7. NUMERICAL RESULTS AND SIMULATIONS

**7.1. Discretization of the evolution equation for  $\chi_t$ .** To investigate the behavior of the pseudo-inverse  $\chi_t$  numerically, we use an implicit finite difference scheme. Following the ideas in [8] we denote the discretized version of  $\chi_t$  by  $\chi_k^i$ , where the spatial discretization is indexed by  $k$  and the temporal discretization by  $i$ . The spatial and temporal step sizes are denoted by  $h$  and  $\tau$ , respectively. A straight forward discretization of the general equation (25) yields

$$(27) \quad \frac{\chi_k^{i+1} - \chi_k^i}{\tau} = - \left( \frac{\kappa(\chi_k^{i+1}, v_f^{i+1})}{(\chi_{k+1}^{i+1} - \chi_k^{i+1})^m} - \frac{\kappa(\chi_{k+1}^{i-1}, v_f^{i+1})}{(\chi_k^{i+1} - \chi_{k-1}^{i+1})^m} \right) h^{m-1} + \lambda(\chi_k^{i+1} - v_f^{i+1}),$$

for  $\eta \in (0, 1)$ , where  $v_f^i = v_f[\chi^i]$ . At the boundary points  $\eta = 0, 1$  the expressions

$$(\chi_k^i - \chi_{k-1}^i)^{-m} \quad \text{and} \quad (\chi_{k+1}^i - \chi_k^i)^{-m}$$

are set to zero, respectively. As stopping criterion for the iteration procedure we use

$$\|\chi^{i+1} - \chi^i\|_{L^2(0,1)} < \text{tol}.$$

Since we expect the density  $\rho_t$  to concentrate at the minimizer  $x_* \in \mathbb{R}^d$  of the cost function  $f$ , the pseudo-inverse  $\chi_t$  should converge towards the constant value  $x_* \in \mathbb{R}^d$ . This causes problems in the computation of the fractions appearing in (27). Our workaround is to evaluate the fractions up to a tolerance and set them artificially to zero if the denominator is too small. The scheme is tested with the well-known benchmark functions for global optimization problems (cf. [3]) shown in Figure 1. The functions (B2) and (B3) may be extended to multi-dimensions, where they are often referred to as the Ackley and Rastrigin function [27] respectively.



**7.2. Particle approximation.** In order to compare the results of the extension  $m > 1$  to the scheme in [31], we are interested in a particle scheme corresponding to the evolution equation for  $m = 2$ . We derive a numerical scheme by rewriting (24) as

$$\partial_t \rho_t = -\partial_x(\mu_t \rho_t) + \partial_{xx}(\kappa_t \rho_t^2) = \partial_x[-\mu_t \rho_t + \rho_t(\partial_x(\kappa_t \rho_t) + \kappa_t \partial_x \rho_t)].$$

The terms  $\partial_x(\kappa_t \rho_t)$  and  $\partial_x \rho_t$  are mollified in the spirit of [25] with the help of a mollifier  $\varphi_\epsilon$ ,

$$\partial_x(\kappa_t \rho_t) \approx \partial_x \varphi_\epsilon * (\kappa_t \rho_t) \quad \text{and} \quad \partial_x \rho_t \approx \partial_x \varphi_\epsilon * \rho_t.$$

Altogether this yields the approximate deterministic microscopic system

$$(28) \quad \dot{X}_t^i = -\lambda(X_t^i - v_t) + \frac{\sigma}{N} \sum_{j=1}^N \partial_x \varphi_\epsilon(X_t^i - X_t^j) \left[ |X_t^j - v_t|^{2p} + |X_t^i - v_t|^{2p} \right],$$

for  $i = 1, \dots, N$ , using the notation of Section 1.

*Remark 7.1.* Note that scheme (28) is deterministic in contrast to the scheme (1) for  $m = 1$ . Unfortunately, it is not trivial to extend the particle scheme for  $m > 2$ .

**7.3. Numerical Results.** In the following, numerical results corresponding to the above discretizations are shown. We use 200 grid points for the spatial discretization of  $\chi_t$  and 500 particles for the particle approximation schemes. Further parameters are fixed as

$$\tau = 2.5 \cdot 10^{-3}, \quad \alpha = 30, \quad \sigma = 0.8, \quad p = 1, \quad \text{tol} = 10^{-6}.$$

The mollifier is chosen to be  $\varphi_\epsilon = \epsilon^{-d} \varphi(x/\epsilon)$ , where

$$\varphi(x) = \frac{1}{Z_d} \begin{cases} \exp\left(-\frac{1}{|x|^2-1}\right), & \text{if } |x| < 1 \\ 0, & \text{else} \end{cases},$$

with normalizing constant  $Z_d$ .

Figure 2 shows the progression of  $\chi_t$  over time corresponding to the benchmarks **(B1)**–**(B3)**. On the left side the case  $m = 1$  is depicted. The tails mentioned in the discussion of (24) can be seen near the boundary point. On the right side the diffusion coefficient is  $m = 2$ , in this case no tails occur as expected.

In [31], the following scheme with an approximate Heaviside function was proposed:

$$(29) \quad dX_t^i = -\lambda(X_t^i - v_t) H_\epsilon(f(X_t^i) - f(v_t)) dt + \sqrt{2}\sigma |X_t^i - v_f| dW_t^i,$$

where  $v_t$  is as given in (1b). Initially, the Heaviside function was added to assure that the particles do not concentrate abruptly. This is essential in cases where the weight parameter  $\alpha > 0$  is chosen too small, thereby yielding a rough approximation of the minimizer at the start of the simulation. In fact, the presence of the Heaviside function prevents particles that attain function values smaller than the function values at the average, i.e.,  $f(X^i) < f(v_t)$ , from drifting in direction of  $v_t$ . In those cases, only the diffusion part is active.

An analogous particle scheme for the porous media equation with  $m = 2$  reads

$$(30) \quad \dot{X}_t^i = -\lambda(X_t^i - v_t) H_\epsilon(f(X_t^i) - f(v_f)) + \frac{\sigma}{N} \sum_{j=1}^N \partial_x \varphi(X_t^i - X_t^j) \left[ |X_t^j - v_t|^{2p} + |X_t^i - v_t|^{2p} \right].$$

For both schemes a smooth approximation of the Heaviside function of the form,

$$H_\epsilon = (1 + \operatorname{erf}(x/\epsilon))/2$$

is used. We therefore compare the results with and without the Heaviside function in Figure 3. In these simulations, we see the damping effect of the Heaviside function. The simulations without Heaviside are faster. Due to the large value of  $\alpha$ , the minimizer is approximated well by  $v_f$ , thus the concentration happens at the actual minimum of the objective functions.

The graphs show the  $L_2$ -distance of  $\mathbf{X}_t$  (left) and  $\chi_t$  (right) to the known minimizer  $x_*$  or equivalently the 2-Wasserstein distance between the solutions of the mean-field equation and the particle scheme to the global consensus at  $\delta_{x_*}$ . The schemes with nonlinear diffusion  $m = 2$  converge faster than their corresponding schemes with linear diffusion. Nevertheless, for practical

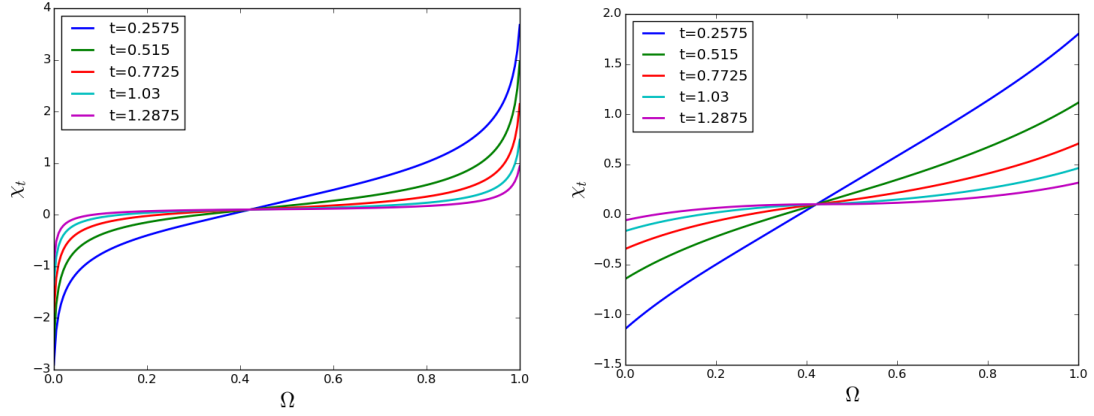
applications with large number of particles, the scheme with linear diffusion is more reasonable due to shorter computation times. Note that in each iteration of the scheme (28) the convolution of all particles has to be computed. The error of the simulation for  $\chi_t$  is smaller than the one for  $\mathbf{X}_t$  at equal times. Even though the benchmarks differ in the steepness of gradients and the number of local minima, the performance of the simulations are comparable. The linear graphs with respect to the logarithmic scaling of the y-axis in Figure 3 indicate the exponential convergence shown in the theoretical section (see Proposition 6.1).

#### ACKNOWLEDGMENTS

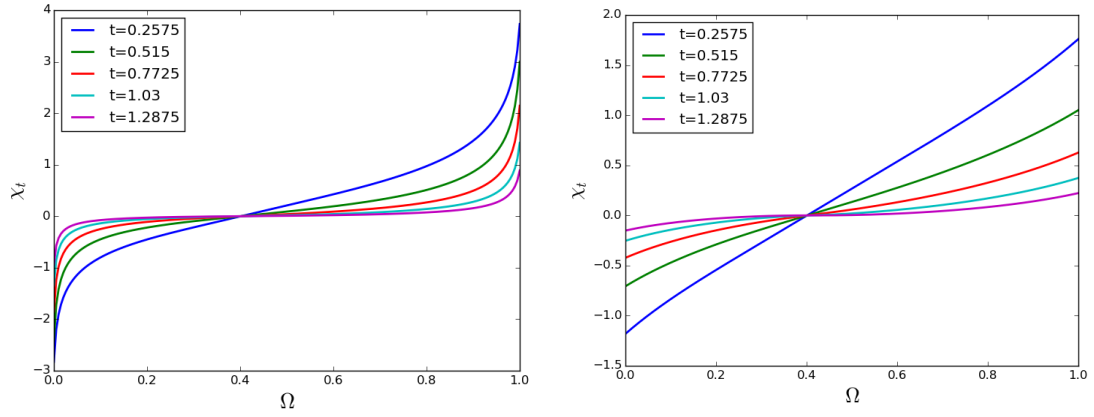
JAC was partially supported by the Royal Society by a Wolfson Research Merit Award. Y-PC is supported by the Alexander Humboldt Foundation through the Humboldt Research Fellowship for Postdoctoral Researchers. CT was partially supported by a 'Kurzstipendium für Doktorandinnen und Doktoranden' by the German Academic Exchange Service.

#### REFERENCES

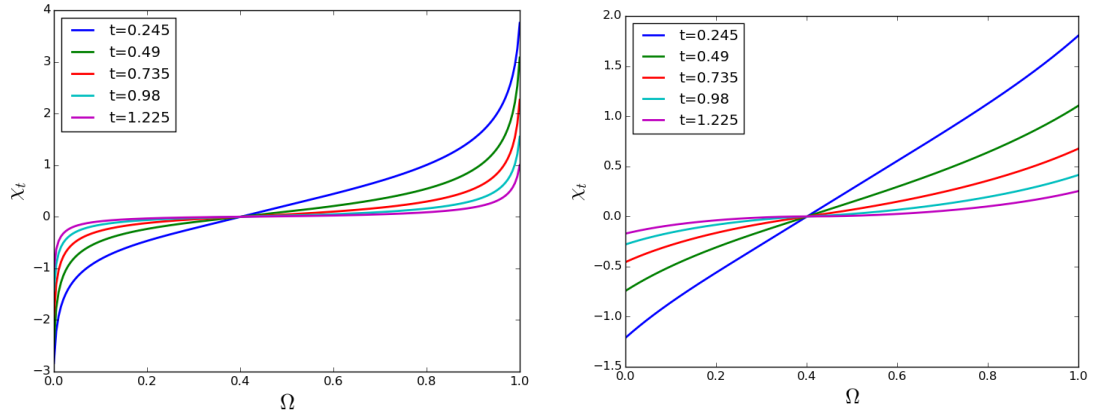
- [1] Emile Aarts and Jan Korst. *Simulated annealing and Boltzmann machines*. New York, NY; John Wiley and Sons Inc., 1988.
- [2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.
- [3] Omid Askari-Sichani and Mahdi Jalili. Large-scale global optimization through consensus of opinions over complex networks. *Complex Adaptive Systems Modeling*, 1(1):1–18, 2013.
- [4] Thomas Back, David B. Fogel, and Zbigniew Michalewicz. *Handbook of evolutionary computation*. IOP Publishing Ltd., 1997.
- [5] Nicola Bellomo, Abdelghani Bellouquid, and Damian Knopoff. From the microscale to collective crowd dynamics. *Multiscale Modeling & Simulation*, 11(3):943–963, 2013.
- [6] Andrea L. Bertozzi, Jesús Rosado, Martin B. Short, and Li Wang. Contagion shocks in one dimension. *J. Stat. Phys.*, 158(3):647–664, 2015.
- [7] Leonora Bianchi, Marco Dorigo, Luca Maria Gambardella, and Walter J Gutjahr. A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing: an international journal*, 8(2):239–287, 2009.
- [8] Adrien Blanchet, Vincent Calvez, and José A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical patlak-keller-segel model. *SIAM Journal on Numerical Analysis*, 46(2):691–721, 2008.
- [9] Christian Blum and Andrea Roli. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR)*, 35(3):268–308, 2003.
- [10] François Bolley. Stochastic mean-field limit: non-lipschitz forces and swarming. *Math. Models Methods Appl. Sci.*, 21(11):2179–2210.
- [11] François Bolley, Ivan Gentil, and Arnaud Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013.
- [12] José A. Carrillo, Young-Pil Choi, and Maxime Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lectures*, pages 1–46. Springer, Vienna, 2014.
- [13] José A. Carrillo, Massimo Fornasier, Jesús Rosado, and Giuseppe Toscani. Asymptotic flocking dynamics for the kinetic cucker-smale model. *SIAM Journal on Mathematical Analysis*, 42(1):218–236, 2010.
- [14] José A. Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, Model. Simul. Sci. Eng. Technol., pages 297–336. Birkhäuser Boston, Inc., Boston, MA, 2010.
- [15] José A. Carrillo, Maria Pia Gualdani, and Giuseppe Toscani. Finite speed of propagation in porous media by mass transportation methods. *Comptes Rendus Mathématique*, 338(10):815–818, 2004.
- [16] José A Carrillo, Yanghong Huang, and Stephan Martin. Explicit flock solutions for quasi-morse potentials. *European Journal of Applied Mathematics*, 25(05):553–578, 2014.
- [17] José A Carrillo, Yanghong Huang, and Stephan Martin. Nonlinear stability of flock solutions in second-order swarming models. *Nonlinear Anal. Real World Appl.*, 17:332–343, 2014.
- [18] Felipe Cucker and Steve Smale. On the mathematics of emergence. *Jpn. J. Math.*, 2(1):197–227, 2007.
- [19] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, 2009.
- [20] Marco Dorigo and Christian Blum. Ant colony optimization theory: A survey. *Theoretical computer science*, 344(2):243–278, 2005.
- [21] Richard Durrett. *Stochastic calculus: a practical introduction*, volume 6. CRC press, 1996.
- [22] Seung-Yeal Ha and Eitan Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinet. Relat. Models*, 1(3):415–435, 2008.



Benchmark (B1)

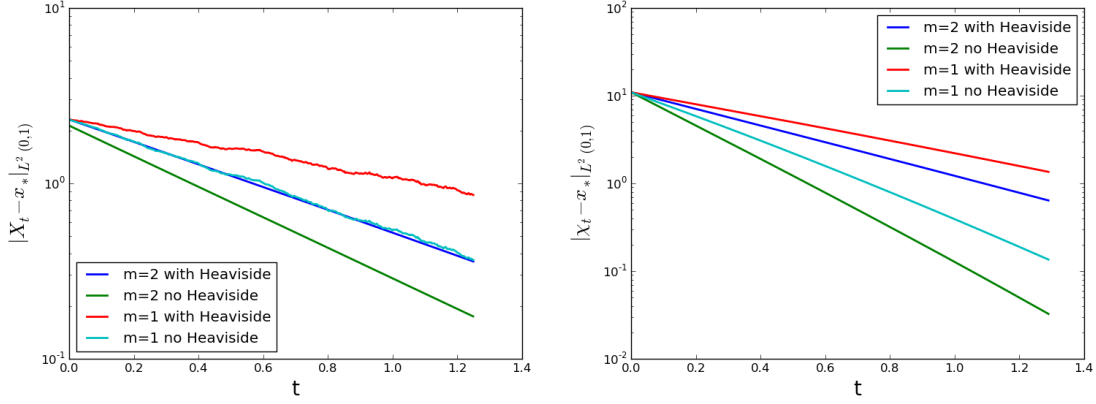


Benchmark (B2)

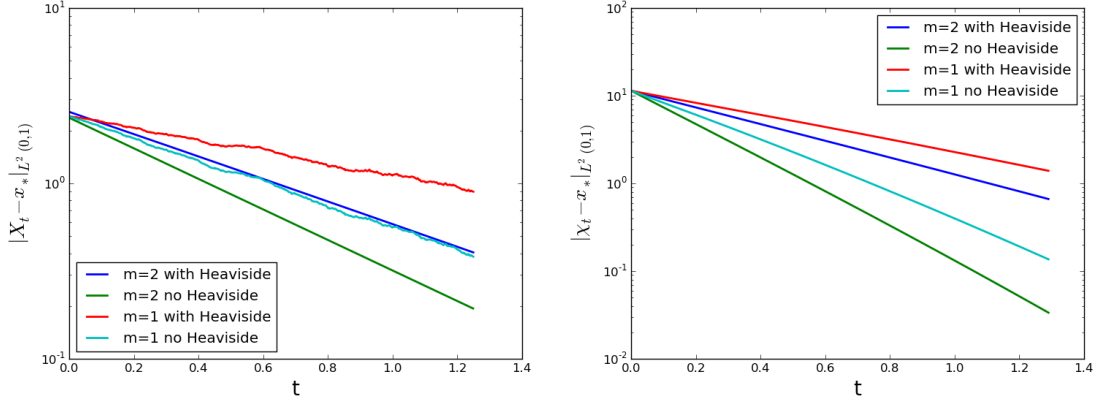


Benchmark (B3)

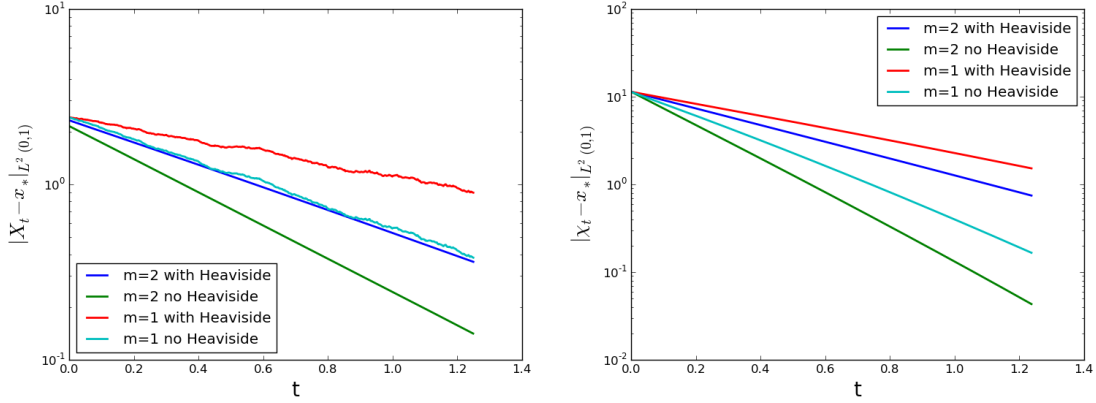
 FIGURE 2. Progression of the inverse distribution function over time for the three benchmarks. Left: Diffusion with  $m=1$ . Right: Diffusion with  $m=2$ .



Benchmark (B1)



Benchmark (B2)



Benchmark (B3)

FIGURE 3.  $L_2$  error of the solution with respect to the minimizer  $x_*$  or, equivalently, the 2-Wasserstein distance between the solution and  $\delta_{x_*}$  for the different benchmarks. Left: Particle scheme. Right: Pseudo-inverse distribution.

- [23] Alfred Horn. Doubly stochastic matrices and the diagonal of a rotation matrix. *American Journal of Mathematics*, 76(3):620–630, 1954.
- [24] James Kennedy. Particle swarm optimization. In *Encyclopedia of Machine Learning*, pages 760–766. Springer, 2010.
- [25] Axel Klar and Sudarshan Tiwari. A multiscale meshfree method for macroscopic approximations of interacting particle systems. *Multiscale Modeling & Simulation*, 12(3):1167–1192, 2014.
- [26] Theodore Kolokolnikov, José A. Carrillo, Andrea Bertozzi, Razvan Fetecau, and Mark Lewis. Emergent behaviour in multi-particle systems with non-local interactions [Editorial]. *Phys. D*, 260:1–4, 2013.
- [27] Xiaodong Li, Ke Tang, Mohammad N Omidvar, Zhenyu Yang, Kai Qin, and Hefei China. Benchmark functions for the cec 2013 special session and competition on large-scale global optimization. *gene*, 7(33):8, 2013.
- [28] Sebastien Motsch and Eitan Tadmor. Heterophilous dynamics enhances consensus. *SIAM Rev.*, 56(4):577–621, 2014.
- [29] Reza Olfati-Saber and Richard M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Automat. Control*, 49(9):1520–1533, 2004.
- [30] Ingram Olkin and Albert W Marshall. *Inequalities: theory of majorization and its applications*, volume 143. Academic press, 2016.
- [31] René Pinnau, Claudia Totzeck, Oliver Tse, and Stephan Martin. A consensus-based model for global optimization and its mean-field limit. *Mathematical Models and Methods in Applied Sciences*, 27(01):183–204, 2017.
- [32] Colin Reeves. *Genetic algorithms*. Springer, 2003.
- [33] Alain-Sol Sznitman. Topics in propagation of chaos. In *Ecole d’été de probabilités de Saint-Flour XIX1989*, pages 165–251. Springer, 1991.
- [34] Giuseppe Toscani. Kinetic models of opinion formation. *Commun. Math. Sci.*, 4(3):481–496, 2006.
- [35] Cédric Villani. Optimal transportation, dissipative pdes and functional inequalities. In *Optimal transportation and applications*, pages 53–89. Springer, 2003.
- [36] Cédric Villani. *Topics in Optimal Transportation (Graduate Studies in Mathematics, Vol. 58)*. American Mathematical Society, 2003.
- [37] James von Brecht, Theodore Kolokolnikov, Andrea L. Bertozzi, and Hui Sun. Swarming on random graphs. *J. Stat. Phys.*, 151(1-2):150–173, 2013.

(José A. Carrillo)

DEPARTMENT OF MATHEMATICS, IMPERIAL COLLEGE LONDON,  
LONDON SW7 2AZ, UNITED KINGDOM

E-mail address: carrillo@imperial.ac.uk

(Young-Pil Choi)

FAKULTÄT MATHEMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN,  
BOLTZMANNSTRASSE 3, D-85748, GARCHING BEI MÜNCHEN, GERMANY

E-mail address: ychoi@ma.tum.de

(Claudia Totzeck)

DEPARTMENT OF MATHEMATICS, TECHNISCHE UNIVERSITÄT KAISERSLAUTERN,  
ERWIN-SCHRÖDINGER-STRASSE, 67663 KAISERSLAUTERN, GERMANY

E-mail address: totzeck@mathematik.uni-kl.de

(Oliver Tse)

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, EINDHOVEN UNIVERSITY OF TECHNOLOGY,  
P.O. BOX 513, 5600MB EINDHOVEN, THE NETHERLANDS

E-mail address: o.t.c.tse@tue.nl